

Non-linear Redshift-Space Power Spectra

J. Richard Shaw^{1,*} and Antony Lewis^{1,†}

¹*Institute of Astronomy, Madingley Road, Cambridge, CB3 0HA, UK.*

(Dated: October 20, 2008)

Distances in cosmology are usually inferred from observed redshifts—an estimate that is dependent on the local peculiar motion—giving a distorted view of the three dimensional structure and affecting basic observables such as the correlation function and power spectrum. We calculate the full non-linear redshift-space power spectrum for Gaussian fields, giving results for both the standard flat sky approximation and the directly-observable angular correlation function and angular power spectrum $C_l(z, z')$. Coupling between large and small scale modes boosts the power on small scales when the perturbations are small. On larger scales power is slightly suppressed by the velocities perturbations on smaller scales. The analysis is general, but we comment specifically on the implications for future high-redshift observations, and show that the non-linear spectrum has significantly more complicated angular structure than in linear theory. We comment on the implications for using the angular structure to separate cosmological and astrophysical components of 21 cm observations.

I. INTRODUCTION

At the most fundamental level cosmological observations consist of measurements of radiation intensity and frequency as a function of angle on the sky. From these we can try to infer properties of the Universe on our past light cone, and from them learn about cosmology. To make more than the simplest inferences we must find a reliable distance to the source we are observing. Fortunately if the frequency of an emitting source is known, the redshift can be used as a measurement of distance, allowing us to map our past light cone as a function of angle and redshift. The observed redshift includes several effects, but the most important is that from cosmological expansion which allows us to estimate the distance. Secondary to this is the doppler shifting from the peculiar velocity of the source along our line of sight. For measurements of the displacement of a source from us, the peculiar velocity quickly becomes negligible in comparison to the cosmological redshifting. However, when measuring the separation between spatially close points the correlated peculiar velocities can have an important effect. When inferring the statistics of cosmological fluctuations it is therefore important to carefully model the effect of velocities.

The universe is assumed to be spatially statistically homogeneous and isotropic at a given time. The non-linear mapping between real space (measured by comoving distance) and redshift space (measured by the redshift z) means that a Gaussian field (with Gaussian densities and velocities) will no longer be Gaussian when observed in redshift space, and its power spectrum will also be different. In this paper we show how to calculate the non-linear redshift-space power spectrum and quantify the effects numerically. The linear result is well-known [1, 2], but here we use a non-perturbative approach to calculate results to all orders. As we shall see, the non-linear corrections can be important at small scales even at high redshift, and are therefore potentially important for future high-redshift observations.

When the non-linear corrections become important, for full consistency one should also calculate the non-linear evolution of the fields: an initially Gaussian random field will be modified once non-linear growth starts to be perturbatively important [3, 4, 5]. These non-linear effects are more complicated to model, and depend on which source is being observed; for example, the 21cm source evolution is quite different to that of galaxy number counts. In this paper we therefore neglect these complications, focussing on understanding the important implications of the redshift-space mapping alone, with the important caveat that our results must be generalized for application to real observations. Our analysis is applicable to any observable that can be reasonably approximated as having a Gaussian source field with Gaussian velocities, and hence, within our approximation, applies equally to biased source number counts or 21cm.

Since the line of sight defines a vector field on the past light cone, the light cone as a function of redshift and angle is only statistically isotropic about the centre of symmetry — the observation point. The inferred angular structure of the field about other points therefore gives information about the local velocity field. In linear theory the velocities are simply related to the total density when dark matter and baryon velocities are the same. Hence an observation

*Electronic address: jrs65@ast.cam.ac.uk

†URL: <http://cosmologist.info>

of the velocities could be used to constrain directly the cosmological density field independently of the sources, which could be hard to model because of complicated astrophysics. We show that the non-linear corrections to the angular structure can be important when attempting to measure the densities this way.

This paper will continue as follows: In the next sub-section (Section I A) we briefly overview the results from linear theory. In Section II we introduce our method for calculating the non-linear redshift-space power spectra. Section III discusses the differences encountered when calculating the power spectrum of radiative fields like the brightness compared to spatial densities such as the matter perturbation. In Section IV we calculate the three-dimensional power spectrum and discuss the results. To go beyond this to the full-sky, in Section V, we calculate the angular correlation function and angular power spectrum. Finally we discuss what bearing our results have on high-redshift 21cm observations in Section VI.

Throughout the rest of this paper we assume a standard flat concordance Λ CDM cosmology with matter densities $\Omega_c h^2 = 0.104$, $\Omega_b h^2 = 0.022$ for dark and baryonic matter respectively. We take a Hubble parameter of $H_0 = 73 \text{ km s}^{-1} \text{ Mpc}^{-1}$, and optical depth to Thomson scattering $\tau = 0.09$. We use a primordial power spectrum with constant spectral index $n_s = 0.95$ and amplitude $A_s = 2.04 \times 10^{-9}$ at a scale of 0.05 Mpc^{-1} . Furthermore we neglect the neutrino masses which should have small effects at high k .

A. Redshift-space mapping and linear result

Assuming the redshift is entirely cosmological, the comoving distance to an object at redshift z is

$$\chi_z = \int_0^z \frac{dz'}{(1+z')\mathcal{H}(z')}, \quad (1)$$

where \mathcal{H} is the comoving Hubble parameter and throughout we use natural units with $c = 1$. In general this equation defines what we call the redshift-space distance, which can easily be calculated from the observed redshift given a background cosmology. However it is not equal to the actual comoving distance in a perturbed universe: the peculiar velocity means that the actual comoving distance χ at redshift z is not χ_z , but also depends on the local velocity field $\mathbf{v}(\mathbf{x})$. Neglecting local evolution of the background, small lensing and general-relativistic effects, and assuming that the peculiar velocities are non-relativistic, the comoving distance is in fact

$$\chi = \chi_z - \mathbf{v}(\mathbf{x}) \cdot \hat{\mathbf{n}} / \mathcal{H}|_z. \quad (2)$$

Note that we assume the peculiar velocity of the observer is removed from the observed redshifts so that only the source velocity matters. From here onwards we write $\phi(\mathbf{x}) \equiv \mathbf{v}(\mathbf{x}) \cdot \hat{\mathbf{n}} / \mathcal{H}$, and denote our coordinates in real space as $\mathbf{x} = \chi \hat{\mathbf{n}}$, and redshift space as $\mathbf{s} = \chi_z \hat{\mathbf{n}}$, such that the mapping between the two is

$$\mathbf{s} = \mathbf{x} + \phi(\mathbf{x}) \hat{\mathbf{n}}. \quad (3)$$

The effect at first order in the power spectrum is well known and easy to calculate [1]. Transforming the mass in a small volume element from real to redshift space using the Jacobian factor we have

$$d^3 s = d^3 x \left| \frac{\partial \mathbf{s}}{\partial \mathbf{x}} \right|. \quad (4)$$

In the distant observer approximation we neglect the curvature of the sky, and thus the Jacobian factor contains only the line of sight term $\frac{\partial s}{\partial \chi} = 1 + \phi'$, with the prime denoting differentiation with respect to the line of sight direction. We discuss this point in more depth in Section III. Conserving the mass in the elements gives

$$\bar{\rho}[1 + \Delta_s(\mathbf{s})] d^3 s = \bar{\rho}[1 + \Delta(\mathbf{x})] d^3 x \quad (5)$$

and hence

$$\Delta_s(\mathbf{s}) = \frac{\Delta(\mathbf{x}) - \phi'(\mathbf{x})}{1 + \phi'(\mathbf{x})}, \quad (6)$$

where the source perturbation in real space is Δ and in redshift space is Δ_s . Expanding this to first order gives the redshift-space perturbation

$$\Delta_s(\mathbf{s}) \approx \Delta(\mathbf{s}) - \phi'(\mathbf{s}). \quad (7)$$

Note that in this we use the fact that $\mathbf{s} = \mathbf{x}$ at first order to transform the arguments. In Fourier space we have

$$\Delta_s(\mathbf{k}) = \Delta(\mathbf{k}) - ik_{\parallel}\phi(\mathbf{k}), \quad (8)$$

where $k_{\parallel} \equiv \hat{\mathbf{n}} \cdot \mathbf{k}$. The quantity we are interested in is the power spectrum P_s of Δ_s given by

$$P_s(\mathbf{k}) = P_{\Delta}(k) + 2ik_{\parallel}P_{\Delta\phi}(\mathbf{k}) + k_{\parallel}^2 P_{\phi}(\mathbf{k}), \quad (9)$$

where P_{Δ} , $P_{\Delta\phi}$ and P_{ϕ} are generated by the obvious contraction of Δ and ϕ . For irrotational flows we can link the velocity vector field to an underlying scalar perturbation δ_v , defined by the relation

$$\nabla \cdot \mathbf{v}(\mathbf{x}) = -\mathcal{H}\delta_v(\mathbf{x}). \quad (10)$$

This definition applies generally and makes no constraints on our tracer Δ . It is, however, motivated by the continuity equation for pressureless matter in the linear growth era. In this regime the scalar perturbation to the velocities δ_v simply relates to the total matter perturbation δ_m via $\delta_v = f\delta_m$, where f is the derivative of the linear growth factor for matter perturbations, $f \equiv d \ln D_+ / d \ln a$. The equivalent Fourier space definition to Eq. (10) is $\mathbf{v}(\mathbf{k}) = i\mathcal{H}(\mathbf{k}/k^2) \delta_v(\mathbf{k})$, so, writing $\mu_k = \hat{\mathbf{n}} \cdot \hat{\mathbf{k}} = k_{\parallel}/k$, Eq. (9) becomes

$$P_s(\mathbf{k}) = P_{\Delta}(k) + 2\mu_k^2 P_{\Delta v}(\mathbf{k}) + \mu_k^4 P_v(\mathbf{k}). \quad (11)$$

If the field we consider is a linearly biased tracer of the underlying matter distribution, such as a simplistic model of galaxy number counts, we would have $\Delta = b\delta_m$, with b the linear bias factor. Assuming no velocity bias this gives the classic Kaiser result (see [1])

$$P_{g,s}(\mathbf{k}) = b^2(1 + b^{-1}f\mu_k^2)^2 P_{\delta}(k). \quad (12)$$

II. NON-LINEAR POWER SPECTRUM

The contributions to the redshift-space power spectrum beyond linear theory could be calculated by a perturbative expansion. We discuss perturbative relationships between real and redshift space in Appendices B and C. However as we might expect this approach becomes tedious above second order, and features independently large terms that nearly exactly cancel. The reason for this behaviour is that the effective displacement caused by a velocity becomes larger than the perturbation wavelength on small scales, so the small scale contribution to $\Delta(\mathbf{s})$ is very different from $\Delta(\mathbf{x})$. However most of this displacement comes from the coherent large-scale velocity field, which has little effect on the *difference* of the velocities that is important for an observable change in the correlation function. A bulk radial displacement is not observable in the flat-sky approximation. For this reason an approach based on transforming the real-space correlation functions may be significantly better. This is the approach we adopt here, which allows us to calculate a simple non-perturbative result for the redshift-space power spectrum.

We would like to find how a Gaussian density field $\Delta(\mathbf{x})$ appears in redshift space. Our starting point is from the conservation of field mass in a small volume element, in both real-space and redshift space

$$[1 + \Delta_s(\mathbf{s})] d^3 s = [1 + \Delta(\mathbf{x})] d^3 x. \quad (13)$$

We emphasize that this is for a density field such as source counts, e.g. the galactic number density. Radiative fields such as the brightness and brightness temperature are different since the measurement is then of observed photon counts, rather than source number counts; we address this in Section III. With this restriction in mind we multiply both sides by $e^{-i\mathbf{k}\cdot\mathbf{s}}$ and integrate, finding that

$$\int [1 + \Delta_s(\mathbf{s})] e^{-i\mathbf{k}\cdot\mathbf{s}} d^3 s = \int [1 + \Delta(\mathbf{x})] e^{-i\mathbf{k}\cdot\mathbf{x}} d^3 x, \quad (14)$$

and substituting $\mathbf{s} = \mathbf{x} + \hat{\mathbf{n}}_x \phi(\mathbf{x})$ we then have

$$(2\pi)^3 \delta^3(\mathbf{k}) + \Delta_s(\mathbf{k}) = \int d^3 x e^{-i\mathbf{k}\cdot\mathbf{x}} [1 + \Delta(\mathbf{x})] e^{-ik_{\parallel}\phi(\mathbf{x})}, \quad (15)$$

where $\delta^3(\mathbf{k})$ is the Dirac delta-function that we can neglect provided we limit ourselves to the behaviour at $\mathbf{k} \neq 0$. To calculate the power spectrum we use

$$\langle \Delta_s(\mathbf{k}) \Delta_s(\mathbf{q}) \rangle = \iint d^3 x d^3 y e^{-i[\mathbf{k}\cdot\mathbf{x} + \mathbf{q}\cdot\mathbf{y}]} \langle [1 + \Delta(\mathbf{x})] [1 + \Delta(\mathbf{y})] e^{-i[k_{\parallel}\phi(\mathbf{x}) + q_{\parallel}\phi(\mathbf{y})]} \rangle, \quad (16)$$

where $q_{\parallel} = \mathbf{q} \cdot \hat{\mathbf{n}}_y$, and $\hat{\mathbf{n}}_y = \mathbf{y}/y$. To calculate the expectation value we assume that all the fields are Gaussian. Writing the fields as a vector $\mathbf{z}^T = (\Delta(\mathbf{x}), \Delta(\mathbf{y}), \phi(\mathbf{x}), \phi(\mathbf{y}))$, and defining a further vector $\mathbf{w}^T = -i(0, 0, k_{\parallel}, q_{\parallel})$, we calculate the expectation values $\langle e^{\mathbf{w}^T \mathbf{z}} \rangle$, $\langle \mathbf{z} e^{\mathbf{w}^T \mathbf{z}} \rangle$ and $\langle \mathbf{z} \mathbf{z}^T e^{\mathbf{w}^T \mathbf{z}} \rangle$, defined by

$$\langle (\dots) e^{\mathbf{w}^T \mathbf{z}} \rangle = \frac{1}{(2\pi)^2 \det^{1/2} \mathbf{C}} \int d^4 z \exp \left[-\frac{1}{2} \mathbf{z}^T \mathbf{C}^{-1} \mathbf{z} + \mathbf{w} \cdot \mathbf{z} \right] (\dots), \quad (17)$$

where the \mathbf{C} is the covariance matrix of the fields $\mathbf{C} = \langle \mathbf{z} \mathbf{z}^T \rangle$. We complete the square in the Gaussian integral to evaluate it, giving

$$\langle e^{\mathbf{w}^T \mathbf{z}} \rangle = e^{\frac{1}{2} \mathbf{w}^T \mathbf{C} \mathbf{w}}. \quad (18a)$$

To calculate the remaining two expectation values we take the partial derivatives with respect to \mathbf{w} :

$$\langle \mathbf{z} e^{\mathbf{w}^T \mathbf{z}} \rangle = e^{\frac{1}{2} \mathbf{w}^T \mathbf{C} \mathbf{w}} \mathbf{C} \mathbf{w}, \quad (18b)$$

$$\langle \mathbf{z} \mathbf{z}^T e^{\mathbf{w}^T \mathbf{z}} \rangle = e^{\frac{1}{2} \mathbf{w}^T \mathbf{C} \mathbf{w}} [\mathbf{C} + \mathbf{C} \mathbf{w} \mathbf{w}^T \mathbf{C}]. \quad (18c)$$

The results of Eq. (18) allow us to evaluate Eq. (16): we take (18a), the 1 and 2 components of (18b), corresponding to $\Delta(\mathbf{x})$ and $\Delta(\mathbf{y})$, and the 1,2 component of (18c), from $\Delta(\mathbf{x})\Delta(\mathbf{y})$, and sum them to construct the expectation value of Eq. (16). The required components are

$$\mathbf{w}^T \mathbf{C} \mathbf{w} = -k_{\parallel}^2 C_{\phi}(\mathbf{x}, \mathbf{x}) - q_{\parallel}^2 C_{\phi}(\mathbf{y}, \mathbf{y}) - 2k_{\parallel} q_{\parallel} C_{\phi}(\mathbf{x}, \mathbf{y}), \quad (19a)$$

$$[\mathbf{C} \cdot \mathbf{w}]_1 + [\mathbf{C} \cdot \mathbf{w}]_2 = -i[q_{\parallel} C_{\Delta\phi}(\mathbf{x}, \mathbf{y}) + k_{\parallel} C_{\Delta\phi}(\mathbf{y}, \mathbf{x})], \quad (19b)$$

$$[\mathbf{C} + \mathbf{C} \mathbf{w} \mathbf{w}^T \mathbf{C}]_{12} = C_{\Delta}(\mathbf{x}, \mathbf{y}) - k_{\parallel} q_{\parallel} C_{\Delta\phi}(\mathbf{x}, \mathbf{y}) C_{\Delta\phi}(\mathbf{y}, \mathbf{x}), \quad (19c)$$

where we have defined $C_{ab}(\mathbf{x}, \mathbf{y}) = \langle a(\mathbf{x})b(\mathbf{y}) \rangle$. Note that statistical isotropy of the underlying correlation requires $\langle \Delta(\mathbf{x})\mathbf{v}(\mathbf{x}) \rangle = 0$ and hence the definition of the ϕ field means that $C_{\Delta\phi}(\mathbf{x}, \mathbf{x}) = 0$. Combining the above, the expectation value $\langle \Delta_s(\mathbf{k}) \Delta_s(\mathbf{q}) \rangle$ evaluates to

$$\begin{aligned} \langle \Delta_s(\mathbf{k}) \Delta_s(\mathbf{q}) \rangle &= \iint d^3 x d^3 y e^{-i[\mathbf{k} \cdot \mathbf{x} + \mathbf{q} \cdot \mathbf{y}]} e^{-\frac{1}{2} [k_{\parallel}^2 C_{\phi}(\mathbf{x}, \mathbf{x}) + q_{\parallel}^2 C_{\phi}(\mathbf{y}, \mathbf{y}) + 2k_{\parallel} q_{\parallel} C_{\phi}(\mathbf{x}, \mathbf{y})]} \\ &\quad \times [1 + C_{\Delta}(\mathbf{x}, \mathbf{y}) - iq_{\parallel} C_{\Delta\phi}(\mathbf{x}, \mathbf{y}) - ik_{\parallel} C_{\Delta\phi}(\mathbf{y}, \mathbf{x}) - k_{\parallel} q_{\parallel} C_{\Delta\phi}(\mathbf{x}, \mathbf{y}) C_{\Delta\phi}(\mathbf{y}, \mathbf{x})]. \end{aligned} \quad (20)$$

This result can now be used to calculate the flat-sky power spectrum $P(\mathbf{k})$ and the directly-observable angular power spectrum $C_l(z, z')$, as we show in the following sections.

It is possible to extend this method to calculation of higher n-point functions, such as the bi-spectrum and higher moments, allowing investigation of the non-Gaussianity introduced solely by the redshift-space distortions. This is conceptually simple, we simply take further moments of Eq. (15) giving

$$\langle \Delta_s(\mathbf{k}_1) \Delta_s(\mathbf{k}_2) \cdots \Delta_s(\mathbf{k}_n) \rangle = \int \left(\prod_{j=1}^n d^3 x_j e^{-i[\mathbf{k}_j \cdot \mathbf{x}_j]} \right) \left\langle \prod_{i=1}^n [1 + \Delta(\mathbf{x}_i)] e^{-i[k_{\parallel i} \phi(\mathbf{x}_i)]} \right\rangle, \quad (21)$$

where we have continued to neglect the behaviour at $\mathbf{k} = 0$. This can be evaluated in the same manner as above, though that is beyond the scope of this paper, we will limit ourselves to the power spectrum.

III. RADIATIVE FIELDS AND THE DISTANT OBSERVER APPROXIMATION

Both the matter density field, and galactic number density are examples of spatial densities where the conserved quantity we consider in the transformation between real and redshift space is the mass in a small volume element

$$\rho_s(\mathbf{s}) d^3 s = \rho(\mathbf{x}) d^3 x. \quad (22)$$

This was the line we proceeded along in the previous section. However for radiative quantities such as the brightness we have a subtly different result: if we radially displace a number of sources we still observe the same number, however

the brightness is less because we receive fewer photons from a source that is more distant. For a detector of area dA , receiving frequencies in a range $d\nu$ about ν from a source region of solid angle $d\Omega$, the brightness I_ν is defined by the energy received dE in a short time dt

$$dE = I_\nu dA d\Omega d\nu dt, \quad (23)$$

or simply the brightness I_ν is the flux onto a detector at a frequency ν from a source per unit solid angle per unit frequency. For radiative fields the fundamental observed quantity is $I_\nu d\Omega d\nu$, the flux in a frequency range ν to $\nu + d\nu$, from a solid angle $d\Omega$. The redshift is determined by the shift from the source frequency ν_0 , and thus the frequency interval $d\nu$ gives the radial distance interval in real or redshift space. The conservation equation, neglecting factors of \mathcal{H} , is then

$$I_\nu(\mathbf{s}) d\Omega ds = I_\nu(\mathbf{x}) d\Omega dx. \quad (24)$$

where the subtle distinction between $I_\nu(\mathbf{s})$ and $I_\nu(\mathbf{x})$ is that in the latter we remove the distortion of the frequency interval $d\nu$ caused by the peculiar motion. In the Rayleigh-Jeans approximation (excellent for typical 21cm line observation) the brightness temperature is $T_b(\nu) \approx I_\nu c^2 / 2k_b \nu^2$, so this result also holds for the brightness temperature. Using $s = x + \phi(\mathbf{x})$ this implies that

$$\Delta_{s,T_b}(\mathbf{s}) = \frac{\Delta_{T_b}(\mathbf{x}) - \phi'(\mathbf{x})}{1 + \phi'(\mathbf{x})}, \quad (25)$$

which was only an approximation in the case of number counts, Eq. (6). We discuss the perturbative expansion of this result in Appendix B. To follow the number count derivation we must take the Fourier transform, and so convert the small parameter space region $d\Omega ds$ into the small volume $d^3s = s^2 d\Omega ds$ (similarly for real space), and hence write Eq. (24) as

$$[1 + \Delta_{s,T_b}(\mathbf{s})] d^3s = [1 + \Delta_{T_b}(\mathbf{x})] \left(1 + \frac{\phi(\mathbf{x})}{x}\right)^2 d^3x. \quad (26)$$

If we simply follow through the analysis of Section II we come unstuck because of the $1 + \phi/x$ term, which would make the analysis significantly more complicated (though not intractable). The simplifying solution is to apply an approximation that is not required in the spatial density case, the distant observer approximation. Given that we are observing at large distances relative to the velocity displacement ϕ , and that the distortions are sourced largely by the gradients of the velocity field, we assert that for all scales of interest $\phi(\mathbf{x})/x \ll \phi'(\mathbf{x})$ and set $1 + \phi/x \approx 1$. At high redshift ($z > 5$) we find ϕ_{rms}/x to be at most of order 10^{-3} whilst ϕ'_{rms} is consistently of order 1, so we expect this to be a reasonable approximation. After this it is possible to apply all the previous analysis to radiative fields such as Δ_{T_b} as well as density fields.

Applying the distant observer approximation not only allows us to consider radiative fields, but allows a simplification of the preceding analysis in all cases. Starting from Eq. (14) we transform the δ -function generating term on the LHS by substituting in explicitly for \mathbf{x} and writing it as

$$\int e^{-i\mathbf{k}\cdot\mathbf{s}} d^3s = \int e^{-i\mathbf{k}\cdot[\mathbf{x} + \hat{\mathbf{n}}\phi(\mathbf{x})]} \left(1 + \frac{\phi(\mathbf{x})}{x}\right)^2 (1 + \phi'(\mathbf{x})) d^3x. \quad (27)$$

Invoking the distant observer approximation removes the ϕ/x term, and canceling off the lowest order terms on both sides leaves us with

$$\Delta_s(\mathbf{k}) = \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}} [\Delta(\mathbf{x}) - \phi'(\mathbf{x})] e^{-ik_{\parallel}\phi(\mathbf{x})}. \quad (28)$$

Note that this holds for all \mathbf{k} including $\mathbf{k} = 0$ unlike the previous formulation. For number counts this approximation neglects first order ϕ/x terms, but for radiative fields it is actually correct to first order, neglecting only terms $\mathcal{O}(\Delta\phi/x)$ and higher. Conceptually this is because if you radially displace a volume at redshift z in an angle $d\Omega$ the physical volume corresponding to that $d\Omega$ increases $\propto r^2$, giving a linear $\mathcal{O}(\phi/x)$ change to the number of sources (c.f. Ref. [6]). However by the inverse-square law the fraction of photons received from each source goes down by $1/r^2$, so the number of photons received is invariant at first order. To ensure that the result for number counts contains all the effects at first order we simply preserve the linear ϕ/x term in Eq. (27).

Comparison with Eq. (25) shows that the quantity $\Delta(\mathbf{x}) - \phi'(\mathbf{x})$ is the source of redshift distortions at first order. Writing the redshift-space perturbation in this form makes it clear where the contributions are coming from, and

more obvious how it reduces to the first order result. Given its importance we will denote the first order source as $\alpha(\mathbf{x}) = \Delta(\mathbf{x}) - \phi'(\mathbf{x})$ from now on. To progress towards the power spectrum we follow the same lines as Eq. (16) to Eq. (20) with the only change that we average over $\mathbf{z}^T = (\alpha(\mathbf{x}), \alpha(\mathbf{y}), \phi(\mathbf{x}), \phi(\mathbf{y}))$ to calculate the expectation. Finally we have the $\langle \Delta_s(\mathbf{k}) \Delta_s(\mathbf{q}) \rangle$ in the distant observer approximation

$$\begin{aligned} \langle \Delta_s(\mathbf{k}) \Delta_s(\mathbf{q}) \rangle = \iint d^3x d^3y e^{-i[\mathbf{k} \cdot \mathbf{x} + \mathbf{q} \cdot \mathbf{y}]} \exp \left(-\frac{1}{2} \left[k_{\parallel}^2 C_{\phi}(\mathbf{x}, \mathbf{x}) + q_{\parallel}^2 C_{\phi}(\mathbf{y}, \mathbf{y}) + 2k_{\parallel} q_{\parallel} C_{\phi}(\mathbf{x}, \mathbf{y}) \right] \right) \\ \times [C_{\alpha}(\mathbf{x}, \mathbf{y}) - k_{\parallel} q_{\parallel} C_{\alpha\phi}(\mathbf{x}, \mathbf{y}) C_{\alpha\phi}(\mathbf{y}, \mathbf{x})] . \end{aligned} \quad (29)$$

To keep this correct for spatial densities at first order we must use $\alpha(\mathbf{x}) = \Delta(\mathbf{x}) - \phi'(\mathbf{x}) - 2\phi(\mathbf{x})/x$, giving the linear result without having assumed the distant observer approximation.

IV. POWER SPECTRA ON THE FLAT-SKY

We first consider the flat-sky approximation, appropriate for a small patch of sky sufficiently thin in redshift that evolution along the light cone can be neglected. The patch is assumed to be at a large distance and subtending a small angle so that $\hat{\mathbf{n}} \approx \hat{\mathbf{n}}'$ across the patch. Since we are neglecting evolution, in a statistically homogenous universe with isotropy broken locally only by the line of sight direction the correlation functions should be a function of $r = |\mathbf{x} - \mathbf{y}|$ and $\mu_r \equiv \hat{\mathbf{n}} \cdot \hat{\mathbf{r}}$ only, so

$$C_{\Delta}(\mathbf{x}, \mathbf{y}) = \xi_{\Delta}(r) \quad (30a)$$

$$C_{\Delta\phi}(\mathbf{x}, \mathbf{y}) = \xi_{\Delta\phi}(r, \mu_r) \quad (30b)$$

$$C_{\phi}(\mathbf{x}, \mathbf{y}) = \xi_{\phi}(r, \mu_r) . \quad (30c)$$

Changing one integration variable from \mathbf{x} to \mathbf{r} in Eq. (20), we can then perform the integration over \mathbf{y} using

$$\int d^3y e^{-i\mathbf{y} \cdot (\mathbf{k} + \mathbf{q})} = (2\pi)^3 \delta^3(\mathbf{k} + \mathbf{q}) . \quad (31)$$

By definition the power spectrum is

$$\langle \Delta_s(\mathbf{k}) \Delta_s(\mathbf{q}) \rangle = (2\pi)^3 \delta^3(\mathbf{k} + \mathbf{q}) P_s(\mathbf{k}), \quad (32)$$

and hence we identify $P_s(\mathbf{k})$ as

$$P_s(\mathbf{k}) = \int d^3r e^{-i\mathbf{k} \cdot \mathbf{r}} \left[1 + \xi_{\Delta}(r) + 2ik_{\parallel} \xi_{\Delta\phi}(r, \mu_r) - k_{\parallel}^2 \xi_{\Delta\phi}(r, \mu_r)^2 \right] e^{-k_{\parallel}^2 [\xi_{\phi}(0) - \xi_{\phi}(r, \mu_r)]} . \quad (33)$$

The correlation functions above are dependent only on the angle between \mathbf{r} and $\hat{\mathbf{n}}$, and hence are azimuthally symmetric, allowing us to integrate out this dependence. If we separate the exponential term as $e^{-i\mathbf{k} \cdot \mathbf{r}} = e^{-ik_{\parallel} r_{\parallel}} e^{-ik_{\perp} r_{\perp} \cos \varphi}$ we can integrate over φ , and use the identity:

$$\frac{1}{2\pi} \int_0^{2\pi} \exp(-i x \cos \varphi) d\varphi \equiv J_0(x), \quad (34)$$

where $J_0(x)$ is the zeroth Bessel function of the first kind. Furthermore knowing that the result will be real, we can write separate real and imaginary parts into the cosine and sine parts of the exponential. Combining these we have

$$\begin{aligned} P_s(\mathbf{k}) = 4\pi \int_0^{\infty} dr \int_0^1 d\mu_r r^2 J_0(k_{\perp} r \sqrt{1 - \mu_r^2}) e^{-k_{\parallel}^2 [\xi_{\phi}(0) - \xi_{\phi}(r, \mu_r)]} \\ \times \left[\cos(k_{\parallel} r \mu_r) \left[1 + \xi_{\Delta}(r) - k_{\parallel}^2 \xi_{\Delta\phi}(r, \mu_r)^2 \right] + 2k_{\parallel} \sin(k_{\parallel} r \mu_r) \xi_{\Delta\phi}(r, \mu_r) \right] . \end{aligned} \quad (35)$$

This is the final form, suitable for numerical evaluation. Unfortunately the integral is highly oscillatory, but we must still include the structure in the integrand across a large range between $k \text{ Mpc} \approx 10^{-3} - 10^3$. This includes a very large number of oscillation and thus requires careful evaluation. Calculation of the correlation functions ξ_{Δ} , $\xi_{\Delta\phi}$ and ξ_{ϕ} from the relevant power spectra is considered in Appendix A.

A result equivalent to Eq. (35) has been derived previously in Ref. [5], but numerical calculation was not attempted because the focus was on low redshifts where other non-linear effects are very important. Here we calculate the effects at high redshift, discuss the physical origin of the various effects, and in Section V also generalize to the directly observable angular power spectrum. We also note from Section III that Eq. (35) can alternatively be written in terms of the correlation functions of the first-order source $\alpha(\mathbf{x})$ as

$$P_s(\mathbf{k}) = 4\pi \int_0^\infty dr \int_0^1 d\mu_r r^2 J_0(k_\perp r \sqrt{1-\mu_r^2}) \cos(k_\parallel r \mu_r) \left[\xi_\alpha(r, \mu_r) - k_\parallel^2 \xi_{\alpha\phi}(r, \mu_r)^2 \right] e^{-k_\parallel^2 [\xi_\phi(0) - \xi_\phi(r, \mu_r)]}. \quad (36)$$

In Figure 1 we compare the non-linear results at redshift 10 for two distinct values of μ_k . There are two distinct effects taking place here: firstly at low μ_k there is a suppression of power across all scales; secondly at high μ_k there is an increase in power which overcomes the general suppression at large values of k . The effect reaches the 1% level at around $k = 0.3h \text{ Mpc}^{-1}$. If we calculate the rms perturbation in spheres of half of this wavelength $\pi/k \approx 10.5h^{-1} \text{ Mpc}$, we find $\sigma_{10.5} \approx 0.077$. Thus at this scale perturbations are still firmly linear, and this effect should be significant relative to any non-linear evolution.

We can gain some insight into the physical origin of these effects by considering the leading order perturbative corrections, that is those second order in the power spectra. We make use of some of the results from Appendix C where we examine the perturbative expansion and second order asymptotics.

The general suppression can be understood from the form of the perturbative result at large scales. Taking the result from (C14), we find that on large scales the non-linear contribution ($\Delta P_s(\mathbf{k}) = P_s(\mathbf{k}) - P_s^{\text{lin}}(\mathbf{k})$) for fully correlated fields is

$$\Delta P_s(\mathbf{k}) \sim -k_\parallel^2 \xi_\phi(0) P_s^{\text{lin}}(\mathbf{k}). \quad (37)$$

To gain insight into this note that $\xi_\phi(0)$ is the point line-of-sight velocity variance in Hubble units, which serves to wash out a large-scale mode with wavenumber \mathbf{k} in the line-of-sight direction by a fraction $\mathcal{O}(k_\parallel \xi_\phi(0)^{1/2})$ of a wavelength. This leads to a suppression of large-scale power.

The expansion of the perturbative result for large k suggests a source of the small-scale boost in power: the superposition of large-scale modes on top of modes at that k . The contributions in Eq. (C11) are complicated, though schematically they are of the form

$$\Delta P_s(\mathbf{k}) \sim P_{\phi'}(\mathbf{k}) \xi_\alpha(0) + P_{\alpha\phi'}(\mathbf{k}) \xi_{\alpha\phi'}(0) + P_\alpha(\mathbf{k}) \xi_{\phi'}(0), \quad (38)$$

where we have neglected constants and angular dependence, and have approximated $k \frac{dP_a(\mathbf{k})}{dk} \approx \text{const.} \times P_a(\mathbf{k})$ which is good for large k in the tail of the spectrum. All terms are of the form power spectrum at some \mathbf{k} times the point variance of another quantity from larger scales. The first term (which is essentially exact) represents the superposition of velocity gradients on the point redshift-space power coming from larger scales. The other terms are similar, but contain complicated angular behaviour which we have omitted.

At lower redshift, when terms above second order become important, the exponent term in Eq. (35) becomes large unless $\mathbf{r} \sim 0$. This leads to an exponential suppression of the coupling from larger scales, reflecting the fact that once small scale velocities effectively wipe out the power by line-of-sight smearing, this wins over the boost due to superimposing larger-scale modes. The calculation is of course not reliable in this regime due to significant non-Gaussianity and non-linear evolution, nonetheless the qualitative effect is well known as the Fingers of God, when non-linear clusters contribute significant small-scale velocities [7]. An extra uncorrelated Gaussian point velocity variance σ_v^2 can easily be included in our model by making the substitution $\xi_\phi(0) \rightarrow \xi_\phi(0) + \sigma_v^2/3\mathcal{H}^2$. This has the effect that $P_s(\mathbf{k}) \rightarrow e^{-k_\parallel^2 \sigma_v^2/3\mathcal{H}^2} P_s(\mathbf{k})$, so that power on scales smaller than the redshift-space spread are exponentially suppressed. This describes the effect of finite line width due to the local thermal motion when considering diffuse 21cm, and also roughly the effect of non-linear virial motion within clusters when measuring number count power spectra on much larger scales. For further discussion of an approximate effective model at low redshift when non-linear evolution is important see Ref. [5, 8].

In Figure 1 we also plot the contributions to the power spectra with the terms at first and second order in the linear spectrum subtracted off, showing the contributions missed by second order perturbation theory. In the $\mu_k = 1$ case these contributions are significant at higher k , being greater than 5% above $k = 10 h^{-1} \text{ Mpc}^{-1}$ —for accurate calculations of the redshift-space power spectrum on small scales a fully non-linear calculation is essential.

In Figure 2 the size of the non-linear contributions from redshift distortions at redshifts of $z = 10$ and $z = 30$ is compared for dark matter and 21cm brightness temperature perturbations. On small scales the boosting of power means that non-linear effects are increasingly important in comparison to the linear prediction. At a redshift of $z = 10$ their dominance at reasonable scales means that they are potentially observationally relevant. This is still true for the 21cm spectra, and we discuss the consequences of this in Section VI.

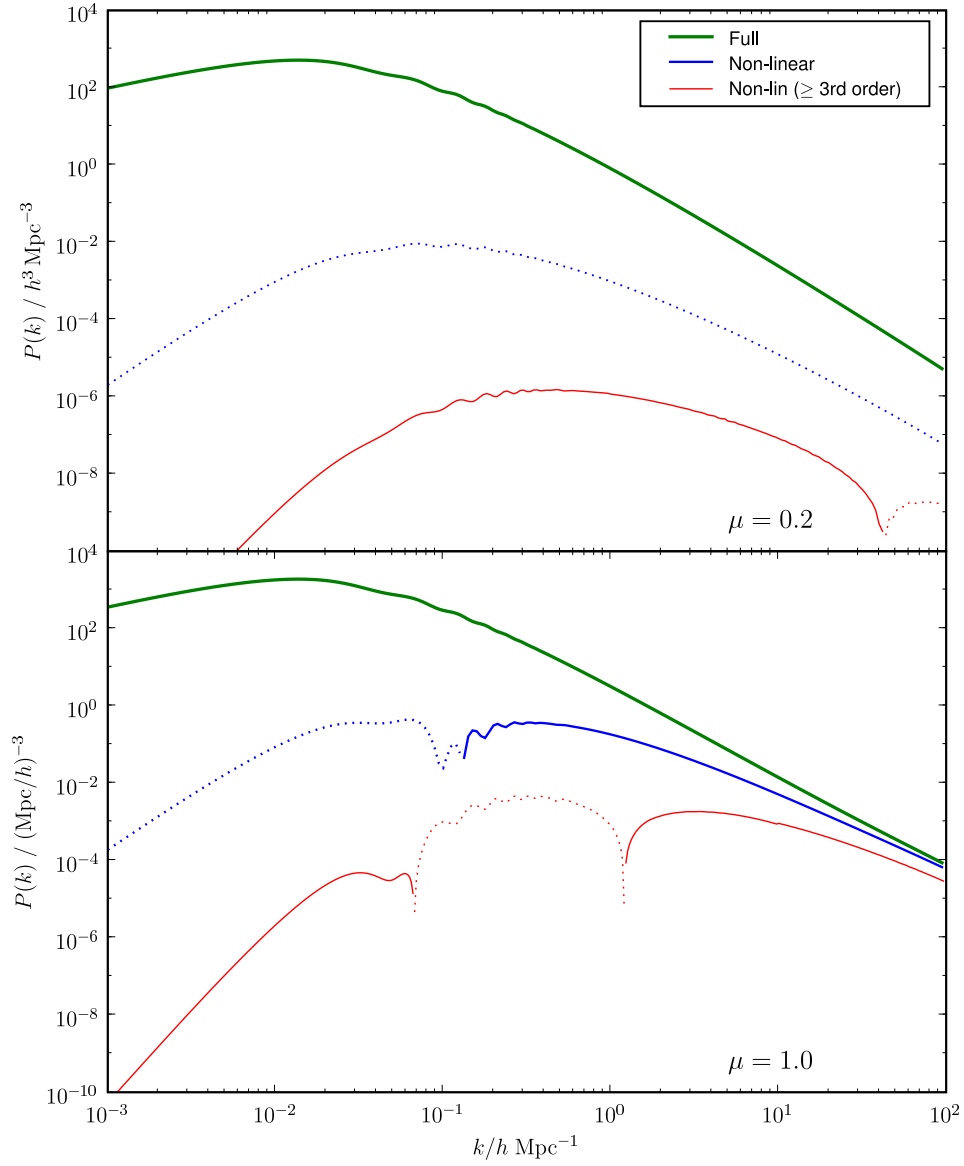


FIG. 1: The full dark matter power spectrum, and the non-linear contributions at redshift $z = 10$. We plot two values of μ , a small value $\mu = 0.2$ in the upper plot and the completely parallel case $\mu = 1.0$ in the lower plot. The solid lines are for positive values, the dotted lines are negative. Whilst the non-linear contributions are negative for $\mu = 0.2$, the contributions at higher μ actually boost the power on small scales. We also plot the non-linear contributions greater than second order in the power spectrum, this shows that second order perturbation theory is largely inadequate at high- k and high μ , at about 5% at $k = 10h \text{ Mpc}^{-1}$.

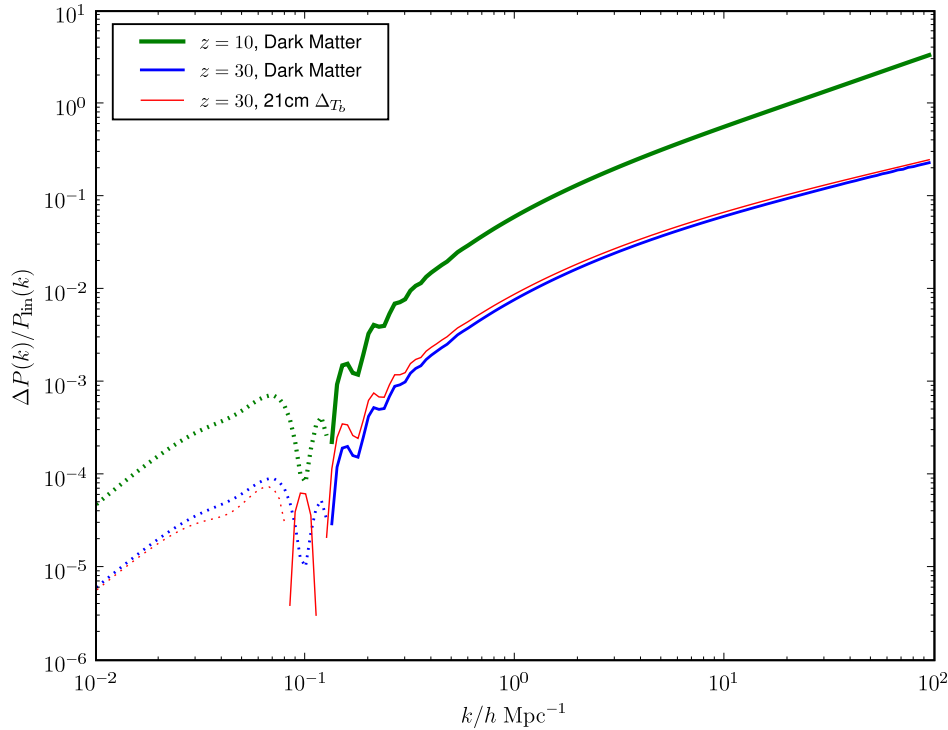


FIG. 2: The ratio of the non-linear contributions to the linear predictions for the dark matter redshift-space power spectrum at redshifts of $z = 10$ and 30 , and the 21cm brightness temperature power spectrum, all for $\mu_k = 1.0$. In all cases the non-linear contributions become significant at high k , whilst in the low redshift dark matter case they become dominant for k approximately greater than 10 hMpc^{-1} . This also shows the 21cm non-linear corrections are of the same magnitude as the dark matter corrections.

In Figure 3 the size of the non-linear contributions from redshift distortions is compared to that from non-linear growth (calculated using 3rd-order perturbation theory [3, 5]). The contributions are of equivalent importance at all scales.

V. ANGULAR CORRELATIONS ON THE CURVED SKY

The redshift space power spectrum that we calculated in the previous section, like the first order result, contains an explicit anisotropy within the small observed volume due to the direction defined by the line of sight. Whilst useful for consideration of localized distortions in redshift space, we should remember that each observer in the universe should see a statistically isotropic light cone if the universe is statistically isotropic and homogeneous. It is the angular correlation between different redshifts on the light cone that is directly observable. The most natural descriptions for the whole sky should take this directly into account, separating out the radial distances and displacements. In this section we calculate the angular correlation function $\xi_s(x, y, \mu)$ which correlates observations at points at redshifts z and z' separated by angle $\cos^{-1} \mu$; and the angular power spectrum $C_l(z, z')$ giving the correlation of multipoles l at different redshifts z and z' .

Our starting point is to calculate the correlation function between positions \mathbf{z} and \mathbf{z}' in redshift space. This is achieved by taking the inverse transform of (20) yielding

$$\langle \Delta_s(\mathbf{z}) \Delta_s(\mathbf{z}') \rangle = \iint \frac{d^3k d^3q}{(2\pi)^6} e^{i[\mathbf{k} \cdot \mathbf{z} + \mathbf{q} \cdot \mathbf{z}']} \langle \Delta_s(\mathbf{k}) \Delta_s(\mathbf{q}) \rangle. \quad (39)$$

This is effectively the forward and inverse transform of our starting point (usually a redundant process), we have required it to eliminate the unwanted k_{\parallel} terms. Substituting (20) into the above (with the delta-function that was

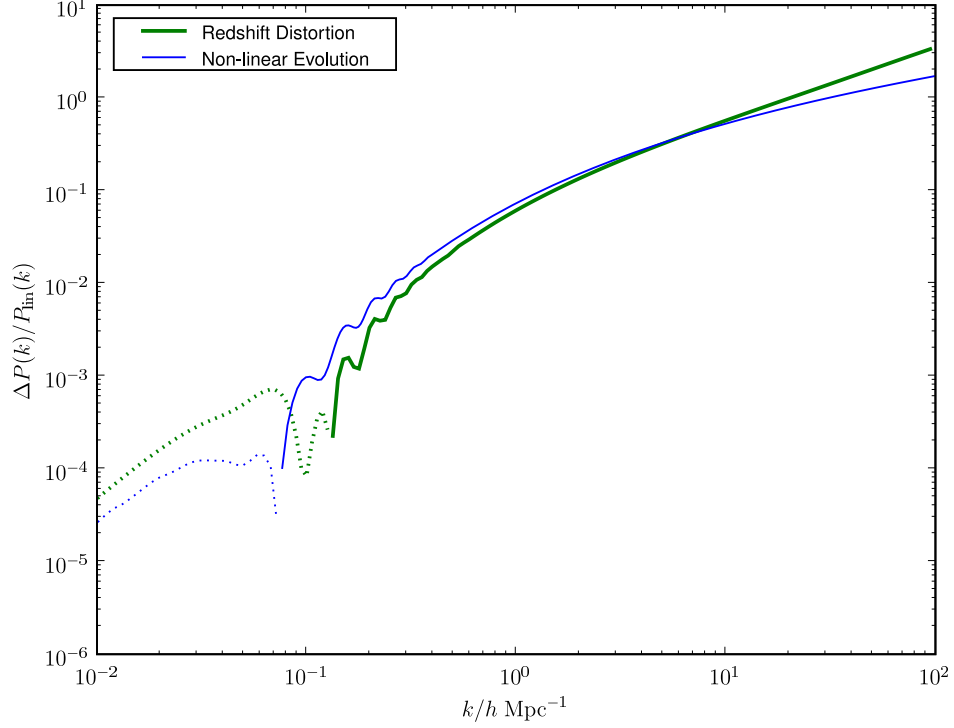


FIG. 3: The ratio of the corrections due to non-linear redshift-space distortions ($\mu_k = 1$) and non-linear evolution contributions to the linear dark matter power spectrum at a redshift of $z = 10$. The redshift distortion corrections are of roughly the same magnitude as those from non-linear growth at all scales, and become greater on smaller scales— both effects should be thought of as equally important when considering modes not orthogonal to the line of sight.

suppressed from Eq. (15)) we have:

$$\begin{aligned} \langle \Delta_s(\mathbf{z}) \Delta_s(\mathbf{z}') \rangle &= \int \frac{d^3k d^3q d^3x d^3y}{(2\pi)^6} e^{i[\mathbf{k} \cdot (\mathbf{z} - \mathbf{x}) + \mathbf{q} \cdot (\mathbf{z}' - \mathbf{y})]} \left[e^{-\frac{1}{2} [k_{\parallel}^2 C_{\phi}(\mathbf{x}, \mathbf{x}) + q_{\parallel}^2 C_{\phi}(\mathbf{y}, \mathbf{y}) + 2k_{\parallel} q_{\parallel} C_{\phi}(\mathbf{x}, \mathbf{y})]} \right. \\ &\quad \times \left. \left[1 + C_{\Delta_s}(\mathbf{x}, \mathbf{y}) - iq_{\parallel} C_{\Delta\phi}(\mathbf{x}, \mathbf{y}) - ik_{\parallel} C_{\Delta\phi}(\mathbf{y}, \mathbf{x}) - k_{\parallel} q_{\parallel} C_{\Delta\phi}(\mathbf{x}, \mathbf{y}) C_{\Delta\phi}(\mathbf{y}, \mathbf{x}) \right] \right] - 1. \end{aligned} \quad (40)$$

The term in the large square brackets above is a function only of $k_{\parallel} = \mathbf{k} \cdot \hat{\mathbf{n}}_x$ and $q_{\parallel} = \mathbf{q} \cdot \hat{\mathbf{n}}_y$, and thus we can integrate out the perpendicular components of \mathbf{k} to give the delta functions $\delta^2(\mathbf{z}_{\perp})$ and $\delta^2(\mathbf{z}'_{\perp})$. These effectively constrain \mathbf{x} and \mathbf{y} enforcing them to be parallel to \mathbf{z} and \mathbf{z}' respectively. Given that redshift distortions displace only along the line of sight this is what we should expect. This leaves the integral:

$$\begin{aligned} \langle \Delta_s(\mathbf{z}) \Delta_s(\mathbf{z}') \rangle &= \int \frac{dk_{\parallel} dq_{\parallel} dx dy}{(2\pi)^2} e^{i[k_{\parallel}(z-x) + q_{\parallel}(z'-y)]} \left[e^{-\frac{1}{2} [k_{\parallel}^2 C_{\phi}(\mathbf{x}, \mathbf{x}) + q_{\parallel}^2 C_{\phi}(\mathbf{y}, \mathbf{y}) + 2k_{\parallel} q_{\parallel} C_{\phi}(\mathbf{x}, \mathbf{y})]} \right. \\ &\quad \times \left. \left[1 + C_{\Delta}(\mathbf{x}, \mathbf{y}) - iq_{\parallel} C_{\Delta\phi}(\mathbf{x}, \mathbf{y}) - ik_{\parallel} C_{\Delta\phi}(\mathbf{y}, \mathbf{x}) - k_{\parallel} q_{\parallel} C_{\Delta\phi}(\mathbf{x}, \mathbf{y}) C_{\Delta\phi}(\mathbf{y}, \mathbf{x}) \right] \right] - 1, \end{aligned} \quad (41)$$

where now the vectors $\mathbf{x} = x\hat{\mathbf{n}}_z$ and $\mathbf{y} = y\hat{\mathbf{n}}'_z$. Conveniently this is now an integral of Gaussian form in the variables k_{\parallel} and q_{\parallel} that we can analytically evaluate. Writing these as the vector $\mathbf{t}^T = (k_{\parallel}, q_{\parallel})$, we recast the integral as

$$\begin{aligned} \langle \Delta_s(\mathbf{z}) \Delta_s(\mathbf{z}') \rangle &= \int \frac{dx dy d^2\mathbf{t}}{(2\pi)^2} \exp \left[-\frac{1}{2} \mathbf{t}^T \mathbf{A}_{\phi} \mathbf{t} - i\mathbf{u}^T \cdot \mathbf{t} \right] \\ &\quad \times \left[1 + C_{\Delta}(\mathbf{x}, \mathbf{y}) - it_2 C_{\Delta\phi}(\mathbf{x}, \mathbf{y}) - it_1 C_{\Delta\phi}(\mathbf{y}, \mathbf{x}) - t_1 t_2 C_{\Delta\phi}(\mathbf{x}, \mathbf{y}) C_{\Delta\phi}(\mathbf{y}, \mathbf{x}) \right] - 1, \end{aligned} \quad (42)$$

where

$$\mathbf{u}^\top = (x - z, y - z') , \quad (43a)$$

$$\mathbf{A}_\phi = \begin{pmatrix} C_\phi(\mathbf{x}, \mathbf{x}) & C_\phi(\mathbf{x}, \mathbf{y}) \\ C_\phi(\mathbf{x}, \mathbf{y}) & C_\phi(\mathbf{y}, \mathbf{y}) \end{pmatrix} . \quad (43b)$$

The prototype for this integral is

$$\int d^2 \mathbf{t} \exp \left(-\frac{1}{2} \mathbf{t}^\top \mathbf{A}_\phi \mathbf{t} - i \mathbf{u}^\top \cdot \mathbf{t} \right) = 2\pi \det^{-1/2} \mathbf{A}_\phi \exp \left(-\frac{1}{2} \mathbf{u}^\top \mathbf{A}_\phi^{-1} \mathbf{u} \right) . \quad (44)$$

Further moments can be generated by taking derivatives with respect to the vector \mathbf{u} as done to construct (18). Putting this together, the correlation function is given by a two dimensional integral in the radial distances x and y ,

$$\begin{aligned} \langle \Delta_s(\mathbf{z}) \Delta_s(\mathbf{z}') \rangle = \int dx dy \frac{e^{-\frac{1}{2} \mathbf{u}^\top \mathbf{A}_\phi^{-1} \mathbf{u}}}{2\pi |\mathbf{A}_\phi|^{1/2}} & \left[1 + C_\Delta(\mathbf{x}, \mathbf{y}) - [\mathbf{A}_\phi^{-1} \mathbf{u}]_2 C_{\Delta\phi}(\mathbf{x}, \mathbf{y}) - [\mathbf{A}_\phi^{-1} \mathbf{u}]_1 C_{\Delta\phi}(\mathbf{y}, \mathbf{x}) \right. \\ & \left. - \left[\mathbf{A}_\phi^{-1} - \mathbf{A}_\phi^{-1} \mathbf{u} \mathbf{u}^\top \mathbf{A}_\phi^{-1} \right]_{12} C_{\Delta\phi}(\mathbf{x}, \mathbf{y}) C_{\Delta\phi}(\mathbf{y}, \mathbf{x}) \right] - 1 . \end{aligned} \quad (45)$$

The result expresses the redshift-space correlation function roughly as the integral of the correlations functions against the Gaussian distribution of the velocities at the two points.

Given the isotropy of the correlation functions $C_a(\mathbf{x}, \mathbf{y})$ they must depend only on the lengths $x = |\mathbf{x}|$, $y = |\mathbf{y}|$ and the angle between them of which we take the cosine $\mu = \hat{\mathbf{n}}_z \cdot \hat{\mathbf{n}}_{z'}$, and so we write them as $C_a(\mathbf{x}, \mathbf{y}) = \xi_a(x, y, \mu)$. Similarly $\langle \Delta_s(\mathbf{z}) \Delta_s(\mathbf{z}') \rangle$ depends only on z , z' and μ , and we write it as $\xi_s(z, z', \mu)$. So in its final form the correlation function is

$$\begin{aligned} \xi_s(z, z', \mu) = \frac{1}{2\pi} \int dx dy \det^{-1/2} \mathbf{A}_\phi \exp \left(-\frac{1}{2} \mathbf{u}^\top \mathbf{A}_\phi^{-1} \mathbf{u} \right) \\ \times \left[1 + \xi_\Delta(x, y, \mu) - [\mathbf{A}_\phi^{-1} \mathbf{u}]_2 \xi_{\Delta\phi}(x, y, \mu) - [\mathbf{A}_\phi^{-1} \mathbf{u}]_1 \xi_{\Delta\phi}(y, x, \mu) \right. \\ \left. + \left[\mathbf{A}_\phi^{-1} - \mathbf{A}_\phi^{-1} \mathbf{u} \mathbf{u}^\top \mathbf{A}_\phi^{-1} \right]_{12} \xi_{\Delta\phi}(x, y, \mu) \xi_{\Delta\phi}(y, x, \mu) \right] - 1 . \end{aligned} \quad (46)$$

This closed form expression completely describes the non-linear redshift-space distortions and unlike the flat sky approach we have yet to make any assumptions about the change along the light cone. This ensures it is easy to incorporate the evolution of the fields and the background [9]. A similar result, specific to the flat-sky was found in [10].

The correlation function is frequently used in the study of baryon acoustic oscillations (BAO) to describe the distortions observed on small patches of sky. There it is conventionally denoted $\xi(\sigma, \pi)$, correlating points separated by a comoving distance along the line-of-sight of π and perpendicular to it σ , where the curvature of the sky is neglected. This gives a total separation $r = \sqrt{\pi^2 + \sigma^2}$, and we place the points an average distance \bar{z} from the origin. We can calculate the non-linear equivalent in the flat-sky by picking z , z' and μ equivalent to σ , π and \bar{z} :

$$z = \sqrt{1 + (\sigma/2\bar{z})^2} (\bar{z} + \pi/2) , \quad (47a)$$

$$z' = \sqrt{1 + (\sigma/2\bar{z})^2} (\bar{z} - \pi/2) , \quad (47b)$$

$$\mu = 2 \tan^{-1} \left(\frac{\sigma}{2\bar{z}} \right) \quad (47c)$$

Figure 4 shows $\xi_s(\sigma, \pi)$ and the difference between the linear and non-linear results, $\Delta\xi_s(\sigma, \pi) = \xi_s(\sigma, \pi) - \xi_\alpha(\sigma, \pi)$, for the exactly parallel and perpendicular cases, calculated by the above procedure. We discuss how to calculate the flat-sky linear correlation function $\xi_\alpha(\sigma, \pi)$ in Appendix A. As in the previous cases the non-linear effects change the correlations on small scales by significant amounts (around 10%), though the effect for the parallel case is much smaller than the perpendicular. In the parallel case there is a smoothing of the acoustic peak, resulting in a small suppression of around 3%.

The distortions introduced on the full sky are perhaps most conveniently described by the the angular correlation function, giving the correlation of multipoles on different redshift slices. The l -th multipole moment $C_l(z, z')$ is found by integrating with $\mathcal{P}_l(\mu)$, the l -th Legendre polynomial, that is

$$C_l(z, z') = 2\pi \int d\mu \mathcal{P}_l(\mu) \xi_s(z, z', \mu) . \quad (48)$$

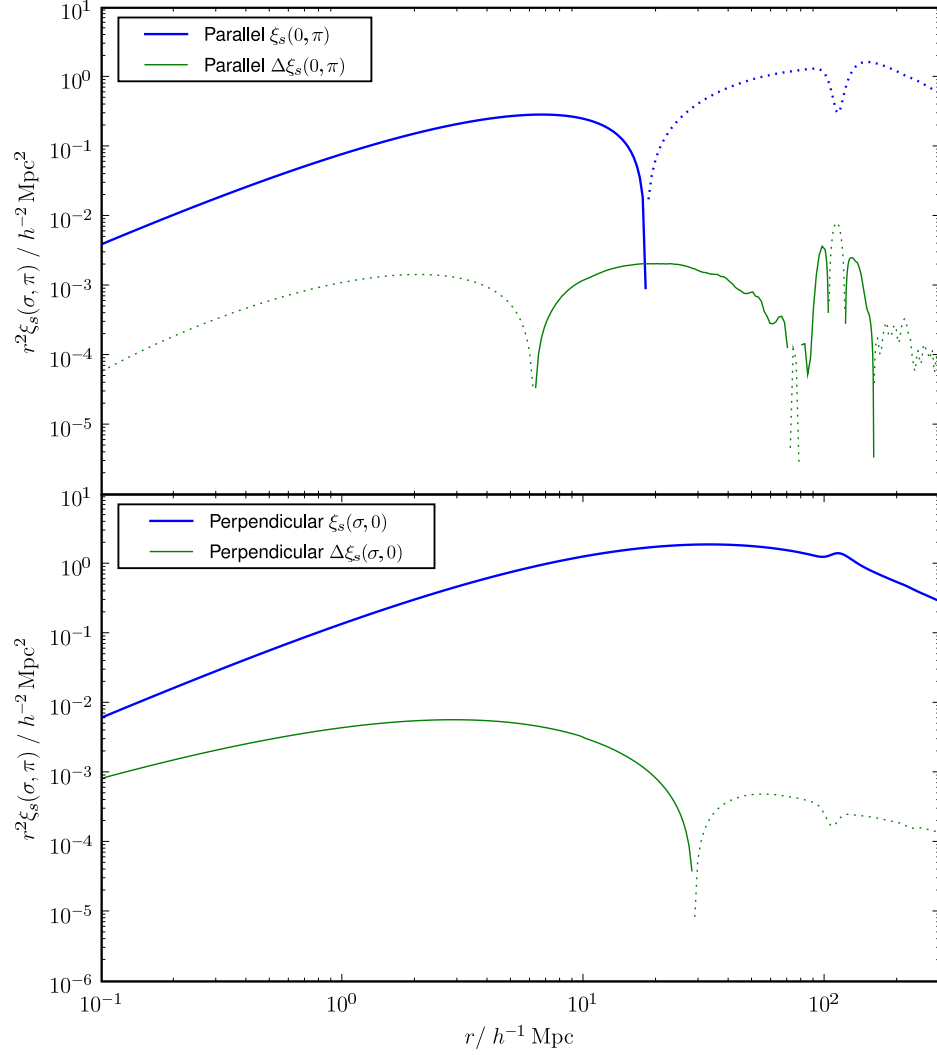


FIG. 4: The redshift-space correlation function $\xi_s(\sigma, \pi)$ at a redshift $z = 10$ for Dark Matter. The top panel illustrates the full correlation function, $\xi_s(\sigma, \pi)$, and the non-linear contributions to it, $\Delta \xi_s(\sigma, \pi)$, in the parallel direction. The lower panel the same, but in the perpendicular direction. The acoustic peak can clearly be seen at a comoving scale of around $100 h^{-1} \text{ Mpc}$. The sharp peaking in the non-linear contributions above $10 h^{-1} \text{ Mpc}$ is largely due to the smoothing effect on the acoustic peak, and small perturbations around the zero crossing points that are large relative to the linear result.

Substituting (45) gives the final integral for the angular correlation (at $l > 0$) for redshifts z and z' :

$$\begin{aligned}
 C_l(z, z') = 2\pi \int d\mu dx dy \mathcal{P}_l(\mu) \frac{1}{2\pi \det^{1/2} \mathbf{A}_\phi} \exp\left(-\frac{1}{2} \mathbf{u}^\top \mathbf{A}_\phi^{-1} \mathbf{u}\right) \\
 \times \left[1 + \xi_\Delta(x, y, \mu) - [\mathbf{A}_\phi^{-1} \mathbf{u}]_2 \xi_{\Delta\phi}(x, y, \mu) - [\mathbf{A}_\phi^{-1} \mathbf{u}]_2 \xi_{\Delta\phi}(y, x, \mu) \right. \\
 \left. + \left[\mathbf{A}_\phi^{-1} - \mathbf{A}_\phi^{-1} \mathbf{u} \mathbf{u}^\top \mathbf{A}_\phi^{-1} \right]_{12} \xi_{\Delta\phi}(x, y, \mu) \xi_{\Delta\phi}(y, x, \mu) \right]. \quad (49)
 \end{aligned}$$

In getting to this result we have avoided most of the common assumptions made when considering redshift-space

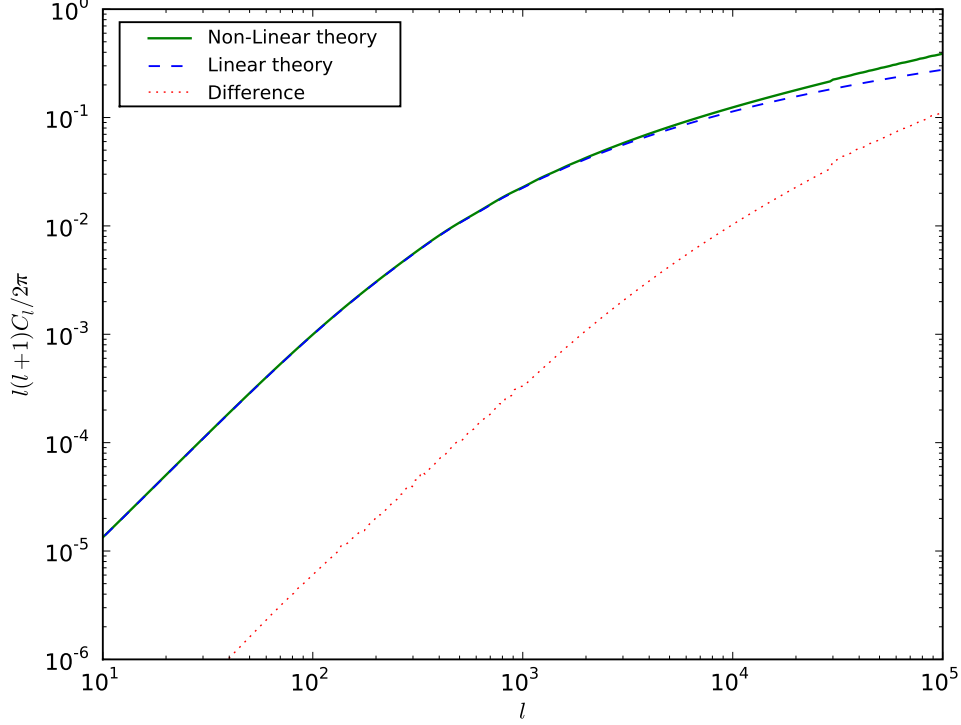


FIG. 5: The equal redshift dark matter angular power spectrum for $z = z' = 10$. We plot the redshift space power spectrum from the linear theory prediction, the non-linear result of Eq. (49) and the difference between the two. The correction is 1% at $l \approx 520$ and becomes greater than 10% above $l \approx 11000$.

problems, non-evolving field statistics and the distant observer approximation (at least for density fields like the matter perturbation, and source number counts). This ensures it naturally incorporates any large angle geometric effects that are not included by taking the flat-sky power spectrum onto the full sky. For further discussion of this see Ref. [6, 11].

The correlation functions $\xi_a(x, y, \mu)$ encapsulate all the information required to calculate the power spectrum, and our formulation above remains completely general. To construct the correlations we must consider several effects, notably the underlying matter correlations and growth along the light cone. In Appendix A we consider how to calculate the correlation functions.

If choosing to use the distant observer approximation, or dealing approximately with radiative fields such as the brightness temperature, we can follow through the same analysis above but starting from the contents of Section III. This leads to the notationally simpler result

$$C_l(z, z') = 2\pi \int d\mu dx dy \frac{e^{-\frac{1}{2}u^T A_\phi^{-1} u}}{2\pi \det^{1/2} A_\phi} \left[\xi_\alpha(x, y, \mu) - \left[A_\phi^{-1} - A_\phi^{-1} u u^T A_\phi^{-1} \right]_{12} \xi_{\alpha\phi}(x, y, \mu) \xi_{\alpha\phi}(y, x, \mu) \right] \mathcal{P}_l(\mu). \quad (50)$$

Note that this is exact at lowest order, only dropping $\mathcal{O}(\phi/x)$ curved-sky terms at higher order, provided we use the correct forms of α for radiative and spatial fields.

In Figure 5 we plot the redshift-space dark matter power spectrum for slices of zero separation at a redshift of $z = 10$, comparing the fully non-linear result to the linear theory (described in detail in [12]). The linear result is essentially the generalisation of the Kaiser result onto the full sky, taking the form

$$C_l(z, z') = \frac{2}{\pi} \int_0^\infty dk k^2 \left[j_l(kz) j_l(kz') P_\Delta(k) - [j_l(kz) j_l''(kz') + j_l''(kz) j_l(kz')] P_{\Delta v}(k) + j_l''(kz) j_l''(kz') P_v(k) \right]. \quad (51)$$

At large l we get a boost in power over the linear-theory results as we would expect from the previous discussion on the flat sky. The effect at small l is less than 1 %, though this is significantly more than the effect on the power

spectrum at equivalent wavenumbers — the lack of intrinsic power at large scales means that the large scale signal in the angular power spectrum is primarily sourced from much higher wave numbers where the non-linear effects are greater. The increases on large scales are a consequence of this with a possible contribution from including the distant observer terms, though we have not disentangled their relative importance. Figure 5 does not obviously show the acoustic peaks, this is a consequence of the fact we do not include a window function in z — the narrow band tends to smooth out such features.

VI. COMPONENT SEPARATION FOR HIGH REDSHIFT 21CM OBSERVATION

The observation of neutral hydrogen through the 21cm spin-flip transition provides a unique opportunity for probing the high-redshift Universe. In principle observations can give a three-dimensional view of structure in the Universe from a redshift of $z = 300$ all the way down to the epoch of reionization at around $z = 6$ and below. The signal seen in absorption at $z \gtrsim 30$ is expected to be nearly linear, with significant redshift distortion [13], and containing angular structure down to the baryon pressure-support scale [14, 15, 16]. With so many modes cosmology could be constrained to very high precision. Although nearly-linear, small non-linear effects will still be very important if observations are to be used reliably, so a non-linear treatment of redshift-distortions will be essential. At redshifts below $z \lesssim 30$ the signal is expected to become much more complicated due to the presence of Lyman- α photons and ionizing sources. Learning about cosmology from these observations would require detailed modelling of complicated and poorly understood astrophysics (see Ref. [17] for a review). Likewise source number counts (in 21cm or otherwise) are hard to model reliably due to scale and time-dependent bias. However in both cases the velocities are likely to be much closer to linear theory, making them a much more robust probe of the underlying cosmological perturbations. If redshift distortions can be isolated, they therefore represent a powerful way to learn about cosmological perturbations from present and near-future observations (e.g. see recent work in Refs. [18, 19] and references therein).

The quantity we are interested in for 21cm observations is the brightness temperature T_b , with perturbation Δ_{T_b} . In real space this is given approximately by

$$\Delta_{T_b} = \beta_b \delta_b + \beta_x \delta_x + \beta_\alpha \delta_\alpha + \beta_{T_K} \delta_{T_K}, \quad (52)$$

where δ_b is the baryon perturbation, δ_x the ionization fraction perturbation, δ_α the Lyman- α coupling perturbation, and δ_{T_K} the perturbation in the gas kinetic temperature. The β_i depend on the background evolution, for a more detailed overview see Ref. [20]. Note that throughout this section we return to the flat-sky approximation.

Although the astrophysics that affects the 21cm signal is very interesting in its own right, to constrain primordial perturbations more directly we would like to determine the power spectrum of matter perturbations $P_\delta(k)$. Unfortunately Δ_{T_b} mixes the astrophysical information from the ionization fraction, Lyman- α coupling and gas temperature in with the cosmological information we desire. However redshift-space distortions add in further information directly linked to the matter perturbations in the approximation in which the source velocities follow the linear CDM velocity. The linear redshift-space power spectrum can then be written

$$P_{s,T_b}(\mathbf{k}) = P_{T_b}(k) + 2\mu_k^2 P_{T_b,v}(k) + \mu_k^4 P_v(k), \quad (53)$$

where the $P_{T_b}(k)$ is the power spectrum of brightness temperature fluctuations in real space encapsulating all the correlations and cross-correlations of Eq. (52). The term $P_{T_b,v}(k)$ gives the cross-correlation with the velocity perturbation δ_v . At linear order we see that the μ_k^4 contribution is entirely the matter power spectrum, giving a possible method of separation without needing to understand the detailed physics encapsulated in $P_{T_b}(k)$ and $P_{T_b,v}(k)$ [21, 22]. However this approach is reliant on the use of the linear expansion: as we can see in Eq. (35) the full angular behaviour is much more complicated, and does not lend itself to an easy separation in powers of μ_k . So we should expect this naive separation method to perform badly wherever the non-linear contributions are important.

To test this in an ideal case, we calculate the theoretical dark matter power spectrum in redshift space at a redshift $z = 10$. Taking 100 points equally spaced in μ_k we integrate with $\mathcal{P}_4(\mu)$, the fourth Legendre polynomial, to isolate the μ_k^4 contribution. With the appropriate normalization, our estimator, exact within linear theory, is

$$\hat{P}_v(k) = \frac{315}{16} \int_{-1}^1 d\mu_k P_s(k, \mu_k) \mathcal{P}_4(\mu_k). \quad (54)$$

We compare the underlying power spectrum with that recovered via this method in Fig. 6. The recovered power spectrum is artificially high at large k . Repeating this with a power spectrum generated from the linear result as expected reproduces the input exactly. In Appendix C we calculate the leading-order non-linear correction on small scales, which shows that we have a direct μ_k^4 contribution taking the form $\mu_k^4 P_v(k) \xi_\alpha(0)$. This combines the

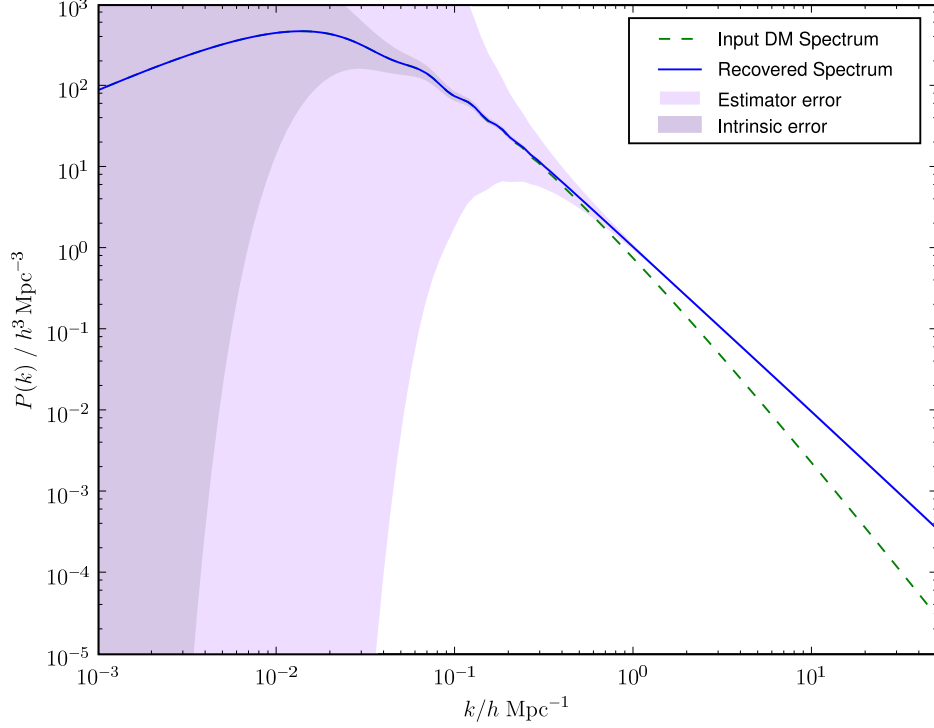


FIG. 6: The input real-space matter power spectrum at $z = 10$ compared to that recovered via the estimator \hat{P}_v given by Eq. (54). We include the errors (shading) for a Hubble volume sized survey at $z = 10$ assuming a binning of $\Delta k/k = 0.1$. The estimator error corresponds to the error if we used the estimator \hat{P}_v discussed in the text. We also plot the intrinsic error that would be seen if we could measure the modes δ_v directly. At high k the recovered spectrum differs dramatically from that input due to the importance of higher-order terms, giving a significant systematic bias outside of the statistical errors.

power spectrum we desire with the source point variance of large scales, mixing in information from the large-scale astrophysics, and is a significant contributor to the bias of this estimator. Correct interpretation of high-redshift observations on small scales will therefore require a more sophisticated analysis that accounts for the complicated angular behaviour introduced at non-linear order, or modelling of the astrophysics in a realistic and accurate manner.

To assess whether any bias is significant, we can calculate the variance of this estimator given a few assumptions about the density of the sampling we can perform in \mathbf{k} -space. We assume a survey of a large volume of the universe V centered at a redshift z , that has a small angular span such that we are still in the flat sky. We define an estimator for the power spectrum at a wavenumber k , and line of sight angle $\cos^{-1} \mu$, that using a suitable weighting function $w_{\mathbf{k}}(k, \mu)$ is defined by

$$\hat{P}_s(k, \mu) = \sum_{\mathbf{k}} w_{\mathbf{k}}(k, \mu) |\Delta_{\mathbf{k}}|^2, \quad (55)$$

where the summation is over all the samples in Fourier space. We are free to choose any weighting function such that the ensemble average $\langle \hat{P}_s(k, \mu) \rangle = P_s(k, \mu)$. Calculating the μ -covariance of this estimator we find

$$\langle \Delta \hat{P}_s(k, \mu_1) \Delta \hat{P}_s(k, \mu_2) \rangle = 2 \sum_{\mathbf{k}} w_{\mathbf{k}}(k, \mu_1) w_{\mathbf{k}}(k, \mu_2) P_s(\mathbf{k})^2. \quad (56)$$

From (54) the variance of the estimator \hat{P}_v is given by

$$\langle \Delta \hat{P}_v(k)^2 \rangle = \frac{99225}{256} \int d\mu_1 d\mu_2 \mathcal{P}_4(\mu_1) \mathcal{P}_4(\mu_2) \langle \Delta \hat{P}_s(k, \mu_1) \Delta \hat{P}_s(k, \mu_2) \rangle \quad (57)$$

Ideally we would optimise the weights $w_{\mathbf{k}}(k, \mu)$ to minimize the variance of \hat{P}_v , but for our purposes it will suffice to pick a representative form—averaging in bins of width Δk and $\Delta \mu$. This picks out $k^2 \Delta k \Delta \mu V / (2\pi)^2 = n(k, \mu) \Delta k \Delta \mu$ modes and we assume that our samples in μ are spaced widely enough that the summation of (56) contributes only when μ_1 equals μ_2 , giving

$$w_{\mathbf{k}}(k, \mu) = \begin{cases} 1/n(k, \mu) \Delta k \Delta \mu & |\mathbf{k}| \in [k, k + \Delta k], \hat{\mathbf{n}} \cdot \mathbf{k} \in [\mu, \mu + \Delta \mu] \\ 0 & \text{otherwise} \end{cases} \quad (58)$$

Given the finite samples in μ we can draw, we approximate the integrals of (57) into summations

$$\langle \Delta \hat{P}_v(k)^2 \rangle \approx \frac{99225}{128} \sum_{ij} (\Delta \mu)^2 \mathcal{P}_4(\mu_i) \mathcal{P}_4(\mu_j) \sum_{\mathbf{k}} w_{\mathbf{k}}(k, \mu_i) w_{\mathbf{k}}(k, \mu_j) P_s(\mathbf{k})^2. \quad (59)$$

Substituting for $w_{\mathbf{k}}(k, \mu)$ connects the summations over i and j , and writing the density of modes with a wavevector of length k as $n(k) = 4\pi k^2 V / (2\pi)^3$ we have

$$\langle \Delta \hat{P}_v(k)^2 \rangle \approx \frac{99225}{64} \frac{1}{n(k) \Delta k} \sum_i (\Delta \mu) \mathcal{P}_4(\mu_i)^2 P_s(k, \mu_i)^2. \quad (60)$$

At low- k the Kaiser result is a reasonable approximation, and thus we use this to calculate the variance. Taking the continuum limit of the summation, we can perform the angular integral analytically for fields with linear bias. The lower bound for the error is the unbiased tracer $b = 1$ giving the numerical result

$$\frac{\Delta \hat{P}_v(k)}{\hat{P}_v(k)} \approx 50 \frac{1}{\sqrt{n(k) \Delta k}}. \quad (61)$$

This shows that the errors in calculating the underlying velocity power spectrum by this component separation are around 35 times larger than those we would find if we could directly measure velocity modes within the observed volume. This increases the lowest k we could infer by around a factor of 10. The plot in Fig. 6 illustrates the dark matter tracing case for which $b = 1$ and the errors are exact. For 21cm we expect to find a large bias and thus the errors are dominated from the contribution of the variance of the $P_{\Delta}(k)$ term. Asymptotically, for large bias

$$\frac{\Delta \hat{P}_v(k)}{\hat{P}_v(k)} \approx 19 \frac{b^2}{\sqrt{n(k) \Delta k}}. \quad (62)$$

To overcome this [23] suggest that combining multiple tracers with distinct biases may be able to reduce this error down closer to the intrinsic level. Though obviously useful for lower redshift surveys where many independent tracers can be found as different galaxy populations, they suggest it may be possible to use this for 21cm observations by applying certain non-linear transformations to the observed field. This method, however, is dependent upon the linear result being correct, restricting its applicability to large scales.

One further ramification is that the higher order angular effects from the non-linear distortions blur any distinction between the Alcock-Paczyński (AP effect) and those of redshift distortions. This may produce complications for methods that seek to obtain cosmological constraints through the AP effect [24]. Generally these provide constraints by tuning parameters until angular dependence of μ^6 and above is eliminated (which is zero for linear redshift-space distortions). However at large k the non-linearities in redshift space ensure that even in the correct cosmology, contributions from higher powers of μ will be non-zero and tuning them to zero would be introducing errors in the parameter fitting. This significance of this is unknown, it may or may not be that the nearly linear low k modes are sufficient to produce constraints unfettered by this.

VII. CONCLUSIONS

We have shown how to calculate the non-linear effects of redshift distortion on the power spectrum in the approximation of Gaussian fields. On small scales the non-linear contributions are important for modes with a component along the line of sight, even at high redshift. Superposition of small-scale power on larger-scale linear modes gives a boost in power on small scales comparable to that from non-linear structure growth. On larger scales smearing by small scale velocities leads to a suppression of power. Any future attempt to extract precision cosmology from high-redshift

observations will need to account carefully for these effects. In order to suitably describe the behaviour on the full-sky we also extended our technique to allow calculation of the angular correlation function and power spectrum. These both have the advantage of naturally incorporating evolution effects of the background and the fields involved, provided they remain Gaussian.

For a fully consistent analysis the non-linear growth and non-Gaussianity should also be accounted for though at present our work does not yet allow this. Despite this we have already demonstrated that just the non-linearities introduced by the mapping from redshift to real space significantly complicate any plan to make accurate measurements of cosmological perturbations by looking for the angular structure in the redshift-space signal from our light cone.

In addition to having a significant effect on the power spectrum and correlation function as discussed in this paper, redshift distortions will also introduce non-Gaussianity. For example there is a non-zero bispectrum for modes that are not all orthogonal to the line of sight. In the approximation of underlying Gaussian fields the method developed in this paper extends straightforwardly to higher n-point functions. This signal will have to be accounted for at high accuracy (along with the bispectrum introduced by non-linear growth) when attempting to use future high-redshift observations to constrain primordial non-Gaussianity [25, 26].

Our work could also be extended to include lensing, which in the Gaussian approximation is just another correlated random field that perturbs points orthogonal to the line of sight.

VIII. ACKNOWLEDGEMENTS

JRS acknowledges receipt of an STFC studentship. AL acknowledges a PPARC/STFC Advanced Fellowship. We thank Anthony Challinor for useful discussion and suggestions.

APPENDIX A: EVALUATING THE CORRELATION FUNCTIONS

To calculate the redshift-space power spectrum we must be able to compute the correlation functions ξ_Δ , $\xi_{\Delta\phi}$ and ξ_ϕ in terms of the matter power spectrum. To start, we note that the 3d-Fourier transform of a radially symmetric function can be simplified dramatically to a 1d transform

$$\int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{r}} f(k) \equiv \frac{1}{2\pi^2} \int_0^\infty dk j_0(kr) [k^2 f(k)], \quad (\text{A1})$$

where $j_0(x) = \sin x/x$ is the zeroth spherical Bessel function. We can generalize this to encapsulate the integrals we will require later on. Expanding in terms of Spherical Harmonics we use the identities for $e^{i\mathbf{k}\cdot\mathbf{r}}$, and $(\hat{\mathbf{n}} \cdot \hat{\mathbf{k}})^n$

$$e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{lm} i^l j_l(kr) Y_{lm}^*(\hat{\mathbf{k}}) Y_{lm}(\hat{\mathbf{r}}), \quad (\text{A2a})$$

$$(\hat{\mathbf{n}} \cdot \hat{\mathbf{k}})^n = 4\pi \sum_{lm} \frac{n!}{(n-l)!!(n+l+1)!!} Y_{lm}^*(\hat{\mathbf{n}}) Y_{lm}(\hat{\mathbf{k}}), \quad (\text{A2b})$$

where $\hat{\mathbf{n}}$ is a direction of our choosing. With these we can easily evaluate integrals of the form

$$\int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{r}} (\hat{\mathbf{k}} \cdot \hat{\mathbf{n}})^n f(k) = \frac{1}{2\pi^2} \sum_{l \leq n} i^l \frac{(2l+1)n!}{(n-l)!!(n+l+1)!!} \mathcal{P}_l(\hat{\mathbf{n}} \cdot \hat{\mathbf{r}}) \int_0^\infty dk [k^2 f(k)] j_l(kr), \quad (\text{A3})$$

where we have used the orthogonality and addition relations of the Spherical Harmonics. For any n the summation only has non-zero elements as far as $l = n$, this means for the small n we are considering the summations will be limited to only a few terms.

Our first assumption is that Δ is a statistically isotropic and homogenous scalar (for example the density perturbation). Secondly we stay with the definition of δ_v from Eq. (10). As a reminder, in real space this relates ϕ and δ_v via

$$\nabla \cdot \mathbf{v}(\mathbf{x}) = -\mathcal{H} \delta_v(\mathbf{x}), \quad (\text{A4})$$

where ∇^{-2} is the inverse Laplacian operator. For observations tracing the underlying matter distribution, $\delta_v = f \delta_m$ exactly in the pressureless limit. We will use the Fourier space equivalent

$$\mathbf{v}(\mathbf{k}) = i\mathcal{H} \frac{\mathbf{k}}{k^2} \delta_v(\mathbf{k}). \quad (\text{A5})$$

We will eventually express the correlations in terms of transforms of the power spectra defined by

$$\langle \Delta(\mathbf{k}; z_x) \Delta(\mathbf{q}; z_y) \rangle = (2\pi)^3 \delta^3(\mathbf{k} + \mathbf{q}) P_\Delta(k; z_x, z_y) , \quad (\text{A6a})$$

$$\langle \Delta(\mathbf{k}; z_x) \delta_v(\mathbf{q}; z_y) \rangle = (2\pi)^3 \delta^3(\mathbf{k} + \mathbf{q}) P_{\Delta v}(k; z_x, z_y) , \quad (\text{A6b})$$

$$\langle \delta_v(\mathbf{k}; z_x) \delta_v(\mathbf{q}; z_y) \rangle = (2\pi)^3 \delta^3(\mathbf{k} + \mathbf{q}) P_v(k; z_x, z_y) , \quad (\text{A6c})$$

which correlate Fourier modes at different epochs given by the redshifts z_x and z_y . In linear theory we can write these in terms of the transfer functions T and the primordial power spectrum P_χ

$$P_\Delta(k; z_x, z_y) = T_\Delta(z_x, k) T_\Delta(z_y, k) P_\chi(k) , \quad (\text{A7a})$$

$$P_{\Delta v}(k; z_x, z_y) = T_\Delta(z_x, k) T_v(z_y, k) P_\chi(k) , \quad (\text{A7b})$$

$$P_v(k; z_x, z_y) = T_v(z_x, k) T_v(z_y, k) P_\chi(k) . \quad (\text{A7c})$$

Numerical calculation of the power spectra can be done via codes such as CAMB [27], or for 21cm perturbations CAMB Sources [16].

The correlation functions can be written in terms of the correlations of Δ and δ_v . Denoting $\theta(\mathbf{x}) = \nabla^{-2} \delta_v(\mathbf{x})$ for brevity, they are

$$C_\Delta(\mathbf{x}, \mathbf{y}) = \langle \Delta(\mathbf{x}) \Delta(\mathbf{y}) \rangle , \quad (\text{A8a})$$

$$C_{\Delta\phi}(\mathbf{x}, \mathbf{y}) = \hat{y}_i \langle \Delta(\mathbf{x}) v_i(\mathbf{y}) \rangle , \quad (\text{A8b})$$

$$C_\phi(\mathbf{x}, \mathbf{y}) = \hat{x}_i \hat{y}_j \langle v_i(\mathbf{x}) v_j(\mathbf{y}) \rangle . \quad (\text{A8c})$$

This reduces the problem down to calculating $\langle \Delta(\mathbf{x}) v_i(\mathbf{y}) \rangle$ and $\langle v_i(\mathbf{x}) v_j(\mathbf{y}) \rangle$. Given the statistical homogeneity and isotropy, these can be decomposed into an isotropic function of the separation $r = |\mathbf{x} - \mathbf{y}|$ combined with the admissible angular factors constructed from $\hat{\mathbf{r}}$.

$$\langle \Delta(\mathbf{x}) \Delta(\mathbf{y}) \rangle = A(r) \quad (\text{A9a})$$

$$\langle \Delta(\mathbf{x}) v_i(\mathbf{y}) \rangle = \mathcal{H} B(r) \hat{r}_i \quad (\text{A9b})$$

$$\langle v_i(\mathbf{x}) v_j(\mathbf{y}) \rangle = \mathcal{H}^2 [C(r) \delta_{ij} + D(r) \hat{r}_i \hat{r}_j] \quad (\text{A9c})$$

where we add the factors of \mathcal{H} for later convenience. $\langle \Delta(\mathbf{x}) \Delta(\mathbf{y}) \rangle$ is scalar function and is simply the transform of the power spectrum P_Δ

$$A(r) = \frac{1}{2\pi^2} \int_0^\infty dk j_0(kr) k^2 P_\Delta(k; z_x, z_y) , \quad (\text{A10})$$

where we leave the z_x, z_y dependence implicit. The other correlation functions are more complicated. There is only one possible direction the vector correlation function $\langle \Delta(\mathbf{x}) v_i(\mathbf{y}) \rangle$ can lie along, the separation vector \mathbf{r} . Multiplying by another \hat{r}_j and contracting, we explicitly find $B(r)$ by substituting substituting the Fourier transform and relating this to the cross power spectrum of Δ and δ_v

$$\begin{aligned} B(r) &= \langle \Delta(\mathbf{x}) \hat{\mathbf{r}} \cdot \mathbf{v}(\mathbf{y}) \rangle / \mathcal{H} \\ &= \int \frac{d^3 k}{(2\pi)^3} e^{i\mathbf{k} \cdot \mathbf{r}} \frac{i\mathbf{k} \cdot \hat{\mathbf{r}}}{k^2} P_{\Delta v}(k; z_x, z_y) \\ &= -\frac{1}{2\pi^2} \int_0^\infty dk j_1(kr) k P_{\Delta v}(k; z_x, z_y) . \end{aligned} \quad (\text{A11})$$

Then correlation of $\langle v_i(\mathbf{x}) v_j(\mathbf{y}) \rangle$ forms a rank-2 tensor that we separate into an isotropic part $C(r)$ and the traceless outer product of \hat{r}_i and \hat{r}_j given by $D(r)$. Taking the trace isolates $C(r)$ and along the same lines as above we find

$$\begin{aligned} C(r) &= \frac{1}{3} \langle \mathbf{v}(\mathbf{x}) \cdot \mathbf{v}(\mathbf{y}) \rangle / \mathcal{H}^2 \\ &= \frac{1}{3} \int \frac{d^3 k}{(2\pi)^3} e^{i\mathbf{k} \cdot \mathbf{r}} \frac{1}{k^2} P_v(k; z_x, z_y) \\ &= \frac{1}{3} \frac{1}{2\pi^2} \int_0^\infty dk j_0(kr) P_v(k; z_x, z_y) . \end{aligned} \quad (\text{A12})$$

Finally we calculate the traceless part $D(r)$

$$\begin{aligned}
D(r) &= \frac{3}{2} \langle v_i(\mathbf{x}) v_j(\mathbf{y}) \rangle (\hat{r}_i \hat{r}_j - \frac{1}{3} \delta_{ij}) / \mathcal{H}^2 \\
&= \frac{3}{2} \int \frac{d^3 k}{(2\pi)^3} e^{i\mathbf{k} \cdot \mathbf{r}} \frac{1}{k^2} P_v(k; z_x, z_y) \left[(\hat{\mathbf{k}} \cdot \hat{\mathbf{r}})^2 - \frac{1}{3} \right] \\
&= -\frac{1}{2\pi^2} \int_0^\infty dk j_2(kr) P_v(k; z_x, z_y) .
\end{aligned} \tag{A13}$$

With these functions calculated we can now express the correlation functions in terms of them

$$C_\Delta(\mathbf{x}, \mathbf{y}) = A(r) , \tag{A14a}$$

$$C_{\Delta\phi}(\mathbf{x}, \mathbf{y}) = \mu_y B(r) , \tag{A14b}$$

$$C_\phi(\mathbf{x}, \mathbf{y}) = \left[C(r) \mu_{xy} + D(r) (\mu_x \mu_y - \frac{1}{3} \mu_{xy}) \right] . \tag{A14c}$$

These results are general, to neaten up the notation somewhat we specialize them to the flat and curved sky cases we have considered. For the flat-sky $\hat{\mathbf{x}} = \hat{\mathbf{y}} = \hat{\mathbf{n}}$, and so $\mu_x = \mu_y = \mu_r$ and $\mu_{xy} = 1$. Evolution along the light cone is also neglected so we evaluate the power spectra at a single fixed redshift z giving

$$\xi_\Delta(r) = A(r) , \tag{A15a}$$

$$\xi_{\Delta\phi}(r, \mu_r) = \mu_r B(r) , \tag{A15b}$$

$$\xi_\phi(r, \mu_r) = \left[C(r) - \frac{1}{3} D(r) \right] + \mu_r^2 D(r) . \tag{A15c}$$

For the curved sky, the correlation function is dependent only on the radial distances of the points, and the angular separation about the origin $\mu = \mu_{xy}$. In terms of these $r = (x^2 + y^2 - 2xy\mu)^{1/2}$, $\mu_x = (y\mu - x)/r$ and $\mu_y = (y - x\mu)/r$ leaving

$$\xi_\Delta(x, y, \mu) = A(r) , \tag{A16a}$$

$$\xi_{\Delta\phi}(x, y, \mu) = \mu_y B(r) , \tag{A16b}$$

$$\xi_\phi(x, y, \mu) = \mu \left[C(r) - \frac{1}{3} D(r) \right] + \mu_x \mu_y D(r) . \tag{A16c}$$

In order to calculate the flat-sky linear redshift correlation function $\xi_\alpha(r, \mu_r)$, we transform the linear redshift-space power spectrum $P_\alpha(\mathbf{k})$, where as we defined earlier $\alpha = \Delta - \phi'$, the linear perturbation in redshift space. Transforming Eq. (11) term by term, again using Eq. (A3), we end up with the following

$$\xi_\alpha(r, \mu_r) = \left[\xi_\Delta^{(0)}(r) + \frac{2}{3} \xi_{\Delta v}^{(0)}(r) + \frac{1}{5} \xi_v^{(0)}(r) \right] - \left[\frac{4}{3} \xi_{\Delta v}^{(2)}(r) + \frac{4}{7} \xi_v^{(2)}(r) \right] \mathcal{P}_2(\mu_r) + \frac{8}{35} \xi_v^{(4)}(r) \mathcal{P}_4(\mu_r) , \tag{A17}$$

where we have defined the correlation-like functions $\xi_a^{(n)}(r)$ by

$$\xi_a^{(n)}(r) = \frac{1}{2\pi^2} \int_0^\infty dk k^2 P_a(k) j_n(kr) . \tag{A18}$$

To use the standard form $\xi_\alpha(\sigma, \pi)$, we simply set $r = \sqrt{\sigma^2 + \pi^2}$ and $\mu_r = \pi/r$.

APPENDIX B: PERTURBATIVE SERIES EXPANSION

In this Appendix we discuss the perturbative expansion of Eq.(6):

$$\Delta_s(\mathbf{s}) = \frac{\Delta(\mathbf{x}) - \phi'(\mathbf{x})}{1 + \phi'(\mathbf{x})} , \tag{B1}$$

where $\mathbf{x} = \mathbf{s} - \phi(\mathbf{x})$. This equation is exact for radiative fields but uses the distant-observer approximation for number counts. To solve this implicit equation for Δ_s we turn to the Lagrange Reversion Theorem¹ that will give us the result in terms of a series expansion. The theorem states that if we have an implicit definition for $v = x + yf(v)$ then the function $g(v)$ is given by the series

$$g(v) = g(x) + \sum_{k=1}^{\infty} \frac{y^k}{k!} \frac{\partial^{k-1}}{\partial x^{k-1}} (f(x)^k g'(x)) . \quad (\text{B2})$$

To obtain $\Delta_s(\mathbf{s})$ we make the obvious assignments to obtain

$$\Delta_s(\mathbf{s}) = \left. \frac{\Delta - \phi'}{1 + \phi'} \right|_{\mathbf{s}} + \sum_{k=1}^{\infty} \frac{1}{k!} \frac{\partial^{k-1}}{\partial \chi^{k-1}} \left[(-\phi)^k \frac{\partial}{\partial \chi} \left(\frac{\Delta - \phi'}{1 + \phi'} \right) \right] \Big|_{\mathbf{s}} . \quad (\text{B3})$$

Expanding $(1 + \phi')^{-1} = \sum_m (-\phi')^m$ and grouping terms of order $n + 1$ this simplifies to

$$\Delta_s(\mathbf{s}) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial \chi^n} [(\Delta - \phi')\phi^n] \Big|_{\mathbf{s}} . \quad (\text{B4})$$

Perturbative results can be obtained using this series expansion, though the perturbative result for the power spectrum is actually obtained more straightforwardly by expansion of the non-perturbative result as we show in Appendix C. The series result can also be written with un-grouped terms as

$$1 + \Delta_s(\mathbf{s}) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial \chi^n} [(1 + \Delta)\phi^n] \Big|_{\mathbf{s}} . \quad (\text{B5})$$

Fourier-transforming Eq. (B4) we have

$$\Delta_s(\mathbf{k}) = \sum_{n=0}^{\infty} \frac{1}{n!} \int d^3 \mathbf{s} e^{-i\mathbf{k} \cdot \mathbf{s}} \frac{\partial^n}{\partial \chi^n} [(\Delta - \phi')(-\phi)^n] \Big|_{\mathbf{s}} \quad (\text{B6})$$

$$\approx \int d^3 x e^{-i\mathbf{k} \cdot \mathbf{x}} [\Delta(\mathbf{x}) - \phi'(\mathbf{x})] e^{-ik_{\parallel} \phi(\mathbf{x})} , \quad (\text{B7})$$

which recovers Eq. (28) of the main text. In the second line we dropped curved-sky corrections from the radial derivatives of x^2 that arise when integrating by parts, which is consistent at linear but not at higher order.

APPENDIX C: PERTURBATIVE RESULT FOR THE REDSHIFT SPACE POWER SPECTRUM

1. General Expansion

Given that we have a general method for calculating the full non-linear result, a perturbative result is perhaps a little crude, however it provides some insight into the source of the most important non-linear effects. We develop the perturbation series from our result for the flat-sky spectrum in terms of the first order source α ,

$$P_s(\mathbf{k}) = \int d^3 r e^{-i\mathbf{k} \cdot \mathbf{r}} \left[\xi_{\alpha}(\mathbf{r}) - k_{\parallel}^2 \xi_{\alpha\phi}(\mathbf{r})^2 \right] e^{-k_{\parallel}^2 [\xi_{\phi}(0) - \xi_{\phi}(\mathbf{r})]} . \quad (\text{C1})$$

First we expand the exponential

$$P_s(\mathbf{k}) = \int d^3 r e^{-i\mathbf{k} \cdot \mathbf{r}} \left[\xi_{\alpha}(\mathbf{r}) - k_{\parallel}^2 \xi_{\alpha\phi}(\mathbf{r})^2 \right] \sum_n \frac{1}{n!} k_{\parallel}^{2n} (\xi_{\phi}(\mathbf{r}) - \xi_{\phi}(0))^n , \quad (\text{C2})$$

¹ See e.g. http://en.wikipedia.org/wiki/Lagrange_reversion_theorem

and then we re-sum the term in ξ_α such that each term in the overall summation contains contributions from the same order in the correlation functions

$$P_s(\mathbf{k}) = P_\alpha(\mathbf{k}) + \sum_n \frac{k_\parallel^{2(n+1)}}{(n+1)!} \int d^3r e^{-i\mathbf{k}\cdot\mathbf{r}} \left[\xi_\alpha(\mathbf{r})\xi_\phi(\mathbf{r}) - \xi_\alpha(\mathbf{r})\xi_\phi(0) - (n+1)\xi_{\alpha\phi}(\mathbf{r})^2 \right] \left(\xi_\phi(\mathbf{r}) - \xi_\phi(0) \right)^n. \quad (\text{C3})$$

The power spectrum P_α can be written in terms of the power spectra of Δ and δ_v , and similarly for the power spectra of $P_{\alpha\phi}$ and P_ϕ :

$$P_\alpha(\mathbf{k}) = P_\Delta(k) + 2\mu_k^2 P_{\Delta v}(k) + \mu_k^4 P_v(k), \quad (\text{C4a})$$

$$P_{\alpha\phi}(\mathbf{k}) = -i\frac{\mu_k}{k} [P_{\Delta v}(k) + \mu_k^2 P_v(k)], \quad (\text{C4b})$$

$$P_\phi(\mathbf{k}) = \frac{\mu_k^2}{k^2} P_v(k). \quad (\text{C4c})$$

Using the convolution theorem we turn the Fourier transform of the products of correlations into a convolution of the corresponding power spectra, giving

$$P_s(\mathbf{k}) = P_\alpha(\mathbf{k}) + \sum_n \frac{k_\parallel^{2(n+1)}}{(n+1)!} (2\pi)^3 \int \frac{d^3k_0}{(2\pi)^3} \frac{d^3k_1}{(2\pi)^3} \frac{d^3k_2}{(2\pi)^3} \delta^3(\mathbf{k}_0 + \mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}) F_n(\mathbf{k}_2) \\ \times \left[P_\alpha(\mathbf{k}_0)P_\phi(\mathbf{k}_1) - P_\alpha(\mathbf{k}_0)\xi_\phi(0)(2\pi)^3\delta^3(\mathbf{k}_1) - (n+1)P_{\alpha\phi}(\mathbf{k}_0)P_{\alpha\phi}(\mathbf{k}_1) \right], \quad (\text{C5})$$

where $\xi_\phi(0)$ is the mean squared line of sight velocity at a point $\xi_\phi(0) = \frac{1}{3}\frac{1}{H^2}\langle \mathbf{v}^2 \rangle$. The convolution kernel $F_n(\mathbf{k})$ is defined as an n -fold convolution of $P_\phi(\mathbf{k}) - (2\pi)^3\xi_\phi(0)\delta^3(\mathbf{k})$,

$$F_n(\mathbf{k}) = (2\pi)^3 \int \frac{d^3q_1}{(2\pi)^3} \cdots \frac{d^3q_n}{(2\pi)^3} [P_\phi(\mathbf{q}_1) - (2\pi)^3\xi_\phi(0)\delta^3(\mathbf{q}_1)] \cdots [P_\phi(\mathbf{q}_n) - (2\pi)^3\xi_\phi(0)\delta^3(\mathbf{q}_n)] \delta^3(\mathbf{q}_1 + \cdots + \mathbf{q}_n - \mathbf{k}), \quad (\text{C6})$$

or equivalently the Fourier transform of the n -th power of $\xi_\phi(\mathbf{r}) - \xi_\phi(0)$:

$$F_n(\mathbf{k}) = \int d^3r e^{-i\mathbf{k}\cdot\mathbf{r}} \left(\xi_\phi(\mathbf{r}) - \xi_\phi(0) \right)^n. \quad (\text{C7})$$

2. Second Order Power Spectrum and Asymptotic Behaviour

In order to gain some intuition into the non-linear redshift space distortions, we turn to the leading order corrections to the linear theory. Using Eq. (C5) we generate the perturbative results to second order in the power spectrum. The lowest order term is simply

$$^{(1)}P(\mathbf{k}) = P_\alpha(\mathbf{k}), \quad (\text{C8})$$

the linear redshift space power spectrum that we expect. The terms at the next order are

$$^{(2)}P(\mathbf{k}) = k_\parallel^2 \left[-P_\alpha(\mathbf{k})\xi_\phi(0) + (2\pi)^3 \int \frac{d^3k_0}{(2\pi)^3} \frac{d^3k_1}{(2\pi)^3} [P_\alpha(\mathbf{k}_0)P_\phi(\mathbf{k}_1) - P_{\alpha\phi}(\mathbf{k}_0)P_{\alpha\phi}(\mathbf{k}_1)] \delta^3(\mathbf{k}_0 + \mathbf{k}_1 - \mathbf{k}) \right], \quad (\text{C9})$$

where at second order in our expansion $n = 0$ and $F_n(\mathbf{k}) = (2\pi)^3\delta^3(\mathbf{k})$ giving the above. Specializing to the case of the matter power spectrum $\Delta = \delta$, and expanding out in full our result is in agreement with that of Ref. [4] when other non-linear effects are neglected.

To investigate the asymptotic behaviour as \mathbf{k} becomes large compared to the turnover in the power spectrum we Taylor expand the above in this limit. We must be careful to include the contributions from where either $|\mathbf{k}_0|$ or $|\mathbf{k}_1|$ are small, as we expect the integral to be dominated by contributions from around the turnover. In this series expansion the leading order terms in $\xi_\phi(0)$ cancel, leaving the dominant term

$$^{(1)}P(\mathbf{k}) = k_\parallel^2 \int \frac{d^3\mathbf{q}}{(2\pi)^3} \left[P_\phi(\mathbf{k})P_\alpha(\mathbf{q}) + \frac{1}{2}q^a q^b \left(P_\phi(\mathbf{q}) [\nabla_a \nabla_b P_\alpha(\mathbf{k})] + P_\alpha(\mathbf{q}) [\nabla_a \nabla_b P_\phi(\mathbf{k})] \right) - 2P_{\alpha\phi}(\mathbf{q}) [\nabla_a P_{\alpha\phi}(\mathbf{k})] q^a \right]. \quad (\text{C10})$$

The $P_\alpha(\mathbf{q}) [\nabla_a \nabla_b P_\phi(\mathbf{k})]$ term above is suppressed by a factor of $(q/k)^2$ relative to the other terms and so we will drop it from our expansion. Averaging out the angular components of the \mathbf{q} integrals removes the summations over a and b and instead directly connects the \mathbf{k} derivatives with the line of sight direction, giving

$$^{(2)}P(\mathbf{k}) \approx k_\parallel^2 \left[P_\phi(\mathbf{k}) \int \frac{d^3q}{(2\pi)^3} P_\alpha(\mathbf{q}) + 2i\hat{\mathbf{n}} \cdot \nabla_k P_{\alpha\phi}(\mathbf{k}) \int \frac{d^3q}{(2\pi)^3} P_{\alpha\phi'}(\mathbf{q}) + \frac{1}{6} [\nabla_k^2 + 2(\hat{\mathbf{n}} \cdot \nabla_k)^2] P_\alpha(\mathbf{k}) \int \frac{d^3q}{(2\pi)^3} P_{\phi'}(\mathbf{q}) \right]. \quad (\text{C11})$$

Each term is of the form of the power spectrum at \mathbf{k} (+derivatives) multiplied by a point variance coming from larger scales. For example the first term gives

$$k_\parallel^2 P_\phi(\mathbf{k}) \xi_\alpha(0) = \mu_k^4 P_v(k) \xi_\alpha(0), \quad (\text{C12})$$

where the point variance of the first order source is

$$\xi_\alpha(0) = \int \frac{d^3q}{(2\pi)^3} \left[P_\Delta(q) + \frac{2}{3} P_{\Delta v}(q) + \frac{1}{5} P_v(q) \right]. \quad (\text{C13})$$

The other terms are more complicated, and for the approximation to make sense the integral ranges should be restricted to scales with $|\mathbf{q}| < |\mathbf{k}|$. The boost in power on small scales can therefore be thought of as due to the superposition of sources at that scale superimposed on large scale linear modes. There are terms up to the sixth power of μ_k .

The behaviour on large scales again can be understood by examining the behaviour for $k \ll k_0, k_1$. Expanding the integral for small k we have

$$^{(2)}P(\mathbf{k}) \approx k_\parallel^2 \left(\frac{1}{3} \int \frac{d^3q}{(2\pi)^3} \frac{1}{q^2} [P_\Delta(q) P_v(q) - P_{\Delta v}(q)^2] - P_\alpha(\mathbf{k}) \xi_\phi(0) + \dots \right). \quad (\text{C14})$$

The first term vanishes in the case of perfect correlation between the source and the velocities, as is the case with one mode of linear perturbations. In this case the dominant contribution is the suppression due to the point line-of-sight velocity variance coming from smaller scales (given by $\xi_\phi(0)$). In the case where the source and velocities do not correlate on large scales, the integral is non-zero and positive (by the Cauchy-Schwarz inequality), reducing the level of suppression.

-
- [1] N. Kaiser, MNRAS **227**, 1 (1987), ADS.
 - [2] A. J. S. Hamilton, in *The Evolving Universe*, edited by D. Hamilton (1998), vol. 231 of *Astrophysics and Space Science Library*, pp. 185–+, arXiv:astro-ph/9708102.
 - [3] N. Makino, M. Sasaki, and Y. Suto, Phys. Rev. D **46**, 585 (1992), ADS.
 - [4] A. F. Heavens, S. Matarrese, and L. Verde, MNRAS **301**, 797 (1998), arXiv:astro-ph/9808016.
 - [5] R. Scoccimarro, Phys. Rev. **D70**, 083007 (2004), arXiv:astro-ph/0407214.
 - [6] P. Papai and I. Szapudi (2008), arXiv:0802.2940.
 - [7] J. C. Jackson, MNRAS **156**, 1P (1972), ADS.
 - [8] W. J. Percival and M. White (2008), arXiv:0808.0003.
 - [9] R. Barkana and A. Loeb, Mon. Not. Roy. Astron. Soc. Lett. **372**, L43 (2006), arXiv:astro-ph/0512453.
 - [10] S. Bharadwaj, MNRAS **327**, 577 (2001), arXiv:astro-ph/0105320.
 - [11] T. Matsubara, ApJ **535**, 1 (2000), arXiv:astro-ph/9908056.
 - [12] K. K. Datta, T. R. Choudhury, and S. Bharadwaj, MNRAS **378**, 119 (2007), arXiv:astro-ph/0605546.
 - [13] S. Bharadwaj and S. S. Ali, MNRAS **352**, 142 (2004), arXiv:astro-ph/0401206.
 - [14] D. Scott and M. J. Rees, MNRAS **247**, 510 (1990), ADS.
 - [15] A. Loeb and M. Zaldarriaga, Phys. Rev. Lett. **92**, 211301 (2004), arXiv:astro-ph/0312134.
 - [16] A. Lewis and A. Challinor, Phys. Rev. **D76**, 083005 (2007), arXiv:astro-ph/0702600.
 - [17] S. Furlanetto, S. P. Oh, and F. Briggs, Phys. Rept. **433**, 181 (2006), arXiv:astro-ph/0608032.
 - [18] L. Guzzo et al., Nature **451**, 541 (2008), arXiv:0802.1944.
 - [19] Y.-S. Song and W. J. Percival (2008), arXiv:0807.0810.
 - [20] S. R. Furlanetto, S. P. Oh, and F. H. Briggs, Phys. Rept. **433**, 181 (2006), arXiv:astro-ph/0608032.
 - [21] R. Barkana and A. Loeb, ApJ **624**, L65 (2005), arXiv:astro-ph/0409572.
 - [22] Y. Mao, M. Tegmark, M. McQuinn, M. Zaldarriaga, and O. Zahn (2008), arXiv:0802.1710.
 - [23] P. McDonald and U. Seljak (2008), arXiv:0810.0323.
 - [24] R. Barkana, MNRAS **372**, 259 (2006), arXiv:astro-ph/0508341.
 - [25] A. Cooray, Phys. Rev. Lett. **97**, 261301 (2006), arXiv:astro-ph/0610257.
 - [26] A. Pillepich, C. Porciani, and S. Matarrese, ApJ **662**, 1 (2007), arXiv:astro-ph/0611126.
 - [27] A. Lewis, A. Challinor, and A. Lasenby, ApJ **538**, 473 (2000), arXiv:astro-ph/9911177.